マルチエンコーダシステムによる 動画像符号化制御方式

正会員 酒 澤 茂 之[†], 正会員 滝 嶋 康 弘[†],
正会員 和 田 正 裕[†], 正会員 羽 鳥 好 律[‡]

Control Scheme for Moving-Picture Coding using the Multi-Encoder System

Shigeyuki Sakazawa[†], Yasuhiro Takishima[†], Masahiro Wada[†] and Yoshinori Hatori[‡]

Abstract This paper describes a control scheme for the multi-encoder system that transmits multiple video signals under the unified control of each encoder. First, the rate-distortion characteristics of the MC+DCT (MPEG) encoder are analyzed. We show that the rate-distortion characteristics can be represented by a simple power function. Based on this function, a bit-rate distribution rule for multiple encoders and a method of applying this to moving pictures is proposed. A computer simulation using real moving pictures shows that the proposed scheme reduces the total distortion of the system as well as probability of severe distortion which occurs in the case of a "difficult" picture, as compared with the non bit-rate distribution method.

1. ま え が き

近年, 衛星放送や CATV などにおいて映像の多チャンネル化が急速に進んでいる。このため, 放送局間の大容量な映像プログラム伝送への要求が高まっている。

通信ネットワークに関しては, ATM 技術を用いた広帯域 ISDN (B-ISDN) の開発が進められてきており, 現在, 実用化の段階へと来ている。B-ISDN は非常に高速な通信サービスおよび, 映像・音声・テキストなど多種多様なトラフィック特性を持つメディアの同時伝送, すなわちマルチメディアサービスを公衆網サービスとして実現することができる¹⁾。その最初期サービスとしては高速通信サービスが提供されることが予

想され, このため大容量回線を用いた映像プログラムの多チャンネル伝送は, ATM 網を用いることが实际的であろう。また, 一般に大容量の回線ほど情報量あたりの通信コストが安くなることから, 複数の映像プログラムを一括伝送する方式が有利である。

複数映像プログラムの一括伝送には, 多重化効果により伝送効率を向上させられるという特長が知られている。一般に, 映像の持つ情報量は時々刻々と変化しており, 符号化により発生する情報量も時間により変動する。このとき, 複数のエンコーダ出力を観察すると, 映像情報源が異なる場合にはそれぞれの発生情報量が同時に大きな値になる確率は小さい。そこで, 複数のエンコーダ出力を多重することにより情報量の総

キーワード: MPEG, 多重化利得, 符号化制御, マルチエンコーダシステム

1994 年 7 月 7 日, テレビジョン学会技術報告に発表

1995 年 12 月 6 日受付, 1996 年 3 月 22 日最終受付

[†] KDD 研究所 (〒 356 上福岡市大原 2-1-15, TEL 0492-78-7426)

[‡] KDD R&D Laboratories (2-1-15, Ohara, Kamifukuoka-shi, Saitama 356, Japan)

和が平滑化されるという大群化効果を得ることができる。本論文では複数のエンコーダ出力を多重し、複数の映像プログラムを一括伝送するマルチエンコーダシステムを提案する。

通常、この大群化効果は独立に動作するエンコーダの場合、最低でも数十台以上のエンコーダ出力の多重が必要である²⁾が、各エンコーダを外部からの統一的制御の下に動作させることにより少数のエンコーダ多重でも効果が得られることが古閑らにより示されている³⁾。そこで提案されている方式は、各エンコーダの出力バッファの占有量に基づいて、各エンコーダの出力バッファの読出しクロックを操作するものである。ここで、エンコーダは出力バッファの占有量に基づいて量子化ステップサイズを制御する。

本論文では、MC+DCT エンコーダにおける符号化映像品質を客観的に把握しつつ制御を行うことを目的として、各エンコーダでのレート歪み特性に基づいてエンコーダ間でのビットレート配分を行う方式を提案する。そして、本方式による符号化利得の解析から、マルチエンコーダシステムの特性を示す。

2. マルチエンコーダシステム

マルチエンコーダシステム⁴⁾は複数のエンコーダを統一的制御の下で動作させるシステムである。本論文においては、その制御目標を総ビットレート一定の下での歪み総和の最小化とする。したがって、マルチエンコーダシステムの制御における操作対象は、各エンコーダの符号化ビットレートとなる。図1にマルチエンコーダシステムの構成図を示す。構成要素はMC+DCT方式のN台のエンコーダ、集中制御装置、およびエンコーダ出力の多重化装置である。

図1 マルチエンコーダシステムの構成
Configuration of a multi-encoder system.

以下に各構成要素の機能について述べる。

集中制御装置は各エンコーダの動作状態を示すパラメータを入力信号として、制御アルゴリズムに基づいてビットレート配分を計算し、各エンコーダへ出力する。

各エンコーダは集中制御装置から指定されたビットレートで動作するためのレートコントロール機構を備えており、画像信号を入力し、符号化ビットストリームと動作状態パラメータを出力する。本論文にて取り扱う符号化アルゴリズムは、映像に対して最も標準的なMC+DCT方式とする。

多重化装置は各エンコーダの出力を多重化し、通信路に送り出す。例えば通信路としてB-ISDNを考える場合には、ATMの1つのVCに複数のビデオストリームを多重することになるが、本論文においてはエンコーダの出力ビットレートの総和のみに着目し、各符号化ビットストリームの多重・分離に関しては考察の対象としない。

3. MC+DCT エンコーダのレート歪み特性

本章では各エンコーダのレート歪み特性の定式化を目的として、MC+DCT符号化器のビットレート対平均二乗誤差（MSE）特性を表現する関数を求める⁵⁾。

ここではレート歪み特性を求めるにあたり、量子化ステップサイズを媒介変数として解析を行う。まず、量子化ステップサイズ対MSEの関係を調べ、ついで量子化ステップサイズ対ビットレートの関係を調べる。最後にその2つの結果を併せることで目的のレート歪み特性を得る。

3.1 ステップサイズ対MSE関係

エンコーダとしてMPEG-2のTM5⁶⁾を用いて、計算機シミュレーションによりステップサイズ Δ 対MSEの関係を調べた。映像FLOWER GARDENとCHEER LEADERの第1予測符号化フレーム（最初のIフレーム直後のPフレーム）において、動き補償誤差信号のDCT係数に対してステップサイズを変えながら(1,1)係数、(6,6)係数および全係数の平均のMSEを求めた結果を図2に示す（ここで、DCT係数のインデックスは0から始まる）。また、図2において、(1,1)、(6,6)係数についてはべき乗関数で、全係数平均については直線でそれぞれ近似したところ、いずれも良好な結果を得た。ここで、それぞれの近似が成り立つ理由について考察を試みる。

まず、低次係数については、一般化ガウス分布⁷⁾に従うことが報告されており、この線形量子化特性は

(a) FLOWER GARDEN

(b) CHEER LEADER

図2 量子化ステップサイズ対 MSE 関係
Relation of quantization step-size and MSE.

MSE を D_l とすると

$$D_l = c_l \Delta^\lambda, \quad 1 < \lambda < 2 \quad (1)$$

となる。

一方、(6,6) 係数に代表される高次係数は電力が小さいため、ステップサイズが大きくなると切り捨てられるので、 Δ -MSE (D_u) 特性としては図2のように頭打ちになる。よって

$$D_u = c_u \Delta^\mu, \quad 0 < \mu < 1. \quad (2)$$

ここで Δ^λ を $\Delta = \alpha$ で Taylor 展開すると

$$\begin{aligned} \Delta^\lambda = & \left\{ (1-\lambda) + \frac{\lambda(\lambda-1)}{2} - \frac{\lambda(\lambda-1)(\lambda-2)}{6} + \dots \right\} \alpha^\lambda \\ & + \left\{ \lambda - \lambda(\lambda-1) + \frac{\lambda(\lambda-1)(\lambda-2)}{2} - \dots \right\} \alpha^{\lambda-1} \Delta \\ & + \left\{ \frac{\lambda(\lambda-1)}{2} - \frac{\lambda(\lambda-1)(\lambda-2)}{2} + \dots \right\} \alpha^{\lambda-2} \Delta^2 + \dots \end{aligned} \quad (3)$$

となる。 Δ^μ についても同様の結果が得られる。このとき、 $\Delta^\lambda + \Delta^\mu$ を考えると $1 < \lambda < 2$, $0 < \mu < 1$ の条件から Δ の1次以外の項が Δ^λ と Δ^μ で互いに逆符号となり、打ち消し合う傾向となる。よって、粗

図3 予測誤差 DCT 信号の性質
Characteristics of prediction error DCT signal.

い近似として次式を得る。

$$D_l + D_u \sim 2 - (\lambda + \mu) + (\lambda + \mu)\Delta \quad (4)$$

すなわち、全係数平均歪みはステップサイズに対して線形特性に近くなる。

3.2 ステップサイズ対ビットレート関係

ビットレートは主として動きベクトルなどのサイド情報に割り当てられる符号量と DCT 係数に割り当てられる符号量から決まる。サイド情報符号量はステップサイズにはあまり依存せず、数 Mb/s 以上の符号化レートではその割合も小さいので、ここでは DCT 係数の符号量について考える。ここで、DCT 係数の符号量は量子化代表値を表現するためのイベント数と1イベントあたりの平均符号長から決まる。そこで、図3に種々の画像における最初のIフレーム直後のPフレームにおける予測誤差信号のDCT係数絶対値の度数分布を示す。図3より、DCT係数絶対値はべき乗的に減少する確率分布となることがわかる。このことから、ステップサイズが大きくなるにつれて、量子化代表値が0ではない有意係数の個数がべき乗的に減少することがわかる。ところで、有意係数の個数は量子化後のジグザク走査を考えると、(ゼロラン、量子化代表値) からなるイベントの個数に等しい。したがって、ステップサイズによらず1イベントあたりの平均符号長があまり変わらなければ、発生情報量はイベント数、すなわち有意係数個数に比例する。以上の考察により、ステップサイズ対ビットレート関係はほぼべき乗関係となるものと予想される。

図4に実験により得た、最初のIフレーム直後のPフレームにおけるステップサイズ対ビットレート関係を示す。グラフ中に示したべき乗近似曲線は良好な近似を与えており、予想したべき乗関係が実際に成り立

図4 量子化ステップサイズ対ビットレート関係
Relation of quantization step-size and bitrate.

つことがわかる．

3.3 レート歪み特性

ステップサイズ対歪み関係が線形であること，ステップサイズ対ビットレート関係がべき乗に従うことから，ビットレート (R) 対 MSE(D) 関係は a, b をパラメータとして次式にて表現できる．パラメータとして次式にて表現できる．

$$D = 10^b R^a \quad (5)$$

$$\log_{10} D = b + a \log_{10} R \quad (6)$$

図5にビットレート対 MSE 関係の実験値によるグラフを示す．図5(a)は線形軸によるグラフ，図5(b)は両対数軸によるグラフであり，いずれも最初の15フレームにわたって平均した MSE の値を用いている．図5(b)における実験値を表すドットが直線で結ばれることから，(5)，(6) 式による関係式が成り立つことがわかる．

ここで，図5(b)における各直線の傾きが画像によらず，ほぼ等しいことが観察される．したがって直線の傾きである a は符号化アルゴリズムに依存するパラメータで，切片に相当する b は画像に依存するパラメータであると考えられる．そこで，(5)，(6) 式における a を一定値と近似し， b の値を画像ごとに図5(b)から求めた．この新しい近似式に基づいて描いた曲線を図5(a)に併せて示す．図5(a)から，この a を一定とする近似が精度良くレート歪み関係を表現していることがわかる．したがって，以下の章では (5)，(6) 式における a の値は一定として取り扱うこととする．

4. ビットレート配分制御方式

4.1 ビットレート配分式

マルチエンコーダシステムの制御に関して，制御の

(a) Linear scale

(b) Bi-logarithmic scale

図5 ビットレート対 MSE 関係
Relation of bitrate and MSE.

目標は各エンコーダで発生する歪み量 (MSE) の総和を最小にすることで，制御の対象となるのは各エンコーダに割り当てるビットレートであり，制約条件はビットレートの総和を一定に保つことである．これらを数学的に表現すれば

$$\text{制約条件} : R_{total} = \sum_{i=1}^N R_i \quad (7)$$

$$\text{最小化式} : D_{total} = \sum_{i=1}^N D_i = \sum_{i=1}^N 10^{b_i} R_i^a \quad (8)$$

となる．ここで， R_i は各エンコーダの割り当てビットレート， R_{total} は総ビットレート， D_{total} は総歪み， N はエンコーダ台数である．この問題は条件付き最小化問題であるので，Lagrange の未定乗数法⁹⁾により， R_i を求めると

$$R_i = \frac{R_{total}}{10^{b_i/(a-1)} \sum_{i=1}^N 10^{-b_i/(a-1)}} \quad (9)$$

となる。

4.2 ビットレート配分の更新

画像のレート歪み特性は画像のシーンによって変化する。一般に映像は時間方向に非定常であるから、時間が経過するとシーンが移り変わって (9) 式のパラメータ b_i も変化する。そこで、時間経過につれて各エンコーダへの割り当てビットレートを更新しなければならない。

本論文ではパラメータを 1 フレームごとに更新することとする。比較的短い間隔で更新する理由は、更新間隔が長すぎると映像シーンの変化に追従できず、特に符号化の易しいシーンから難しいシーンへ移ったときに重大な画質劣化を引き起こすためである。

更新の方法は、まず直前に符号化したフレームのレート R_i と歪み D_i の測定値を (5) 式へ代入し、固定値 a を用いてパラメータ b_i を決定する。ついで、このようにして各エンコーダで求めたパラメータ値 b_i を (9) 式に代入することで、各々のエンコーダに対する割り当てビットレートを求めることができる。

4.3 マルチエンコーダ利得

本節では、エンコーダ間でビットレート配分を行うことによる利得をマルチエンコーダ利得と定義し、マルチエンコーダシステムの性能の特性について検討する。

マルチエンコーダ利得は、ビットレート配分制御を行う場合の MSE の総和に対する、行わない場合の MSE の総和の比と定義される。制御を行わない場合、各エンコーダに指定されるビットレートは、総ビットレート R_{total} をエンコーダ台数 N で割った値 (R_{total}/N) とする。したがって、制御を行わない場合の MSE 総和 D_{total}^* は (8) 式から直ちに得られる。一方、制御を行う場合の MSE 総和 D_{total} は、(9) 式を (8) 式に代入して得られる。以上より、マルチエンコーダ利得 G は D_{total} と D_{total}^* の比であるから (10) 式となる。

$$G = \frac{\sum_{i=1}^N \frac{10^{b_i}}{N^a}}{\sum_{j=1}^N \frac{10^{b_j}}{10^{b_j/(a-1)} \sum_{k=1}^N 10^{-b_k/(a-1)}}} \quad (10)$$

まず、マルチエンコーダ利得の最大値について、エンコーダ 2 台の場合について解析する。 N が 2 の場合の (10) 式におけるパラメータ b_1, b_2 を変化させてマルチエンコーダ利得 G を数值的に計算したところ、 b_1

図6 マルチエンコーダ利得特性
Characteristics of multi-encoder gain.

を固定して b_2 を大きくすると G の最大値は 2.02 に収束した。ここで a の値は図 5(b) より $a = -1.015$ である。一方、 G は $b_1 = b_2$ の時に最小となり、最小値は 1 となった。すなわち、(9) 式を用いたマルチエンコーダシステムは、エンコーダを制御しない場合と比較して性能が劣ることはない。マルチエンコーダシステムは割り当てビットレートを各画像に対して等しくすることができるので、ビットレート割り当て制御のない場合を含むためである。逆に利得が最大になるのは、 b_1 と b_2 がかなり異なるとき、すなわち符号化の難易が異なる画像がペアになっているときである。一般の場合を考えてみると、エンコーダの台数が多くなるほど利得の最大値は増加する。このとき最大利得を得るための条件は、ひとつの画像のみが難しく、他の画像すべてが易しいことである。なぜなら、易しい画像に割り当てるビットレートを減らして難しい画像へ割り当てるビットレートを増やす際に、易しい画像の数が多いほど難しい画像に割り当てられるビットレートが増えるためである。

ところで、利得が大きくなるのは画像の符号化難易度が極度に異なる場合で、現実の利得の値は適用する画像に依存する。そこで、図 6 に画像 “FLOWER GARDEN” (F), “MOBILE & CALENDAR” (M), “SUSIE” (S) を 2 つずつ組合せた場合 (2 チャンネル多重) の座標点を示す。ここで、 b_i の値は各画像の最初の 15 フレームを用いて求めた値である。図 6 において、利得の最大値は 2.02 であるのに対して、最大利得となる組合せ (S,M) でも 1.40 である。

5. 計算機シミュレーション

ビットレート配分を行う提案方式と、すべてのエンコーダが同じビットレートで動作する従来方式を比較する目的で、平均ビットレート対平均歪み特性、割り当てビットレートの時間変化、そして符号化画像

(a) エンコーダ 3 台の場合

(b) エンコーダ 4 台の場合

図 7 符号化特性
Coding characteristics.

の局所的な符号化歪みの分布について実画像を用いた計算機シミュレーションを行った。シミュレーションに際して、MPEG-2 TM5 をエンコーダとして用いた。TM5 のパラメータは、MP@ML, $N=15$, $M=3$, Dual Prime 予測有りである。

図 7 に符号化特性を示す。これは各画像の最初の 15 フレームに対する平均ビットレートと平均 MSE の関係について、提案方式と従来方式におけるシミュレーション結果および、(10) 式の分母から求めたマルチエンコーダシステムの理論 MSE 値を示している。図 7(a) はエンコーダが 3 台の場合、(b) はエンコーダが 4 台の場合をそれぞれ示す。3 台の場合に用いた画像は“POPPLE”、“MOBILE & CALENDAR”、“SUSIE”で、4 台の場合はそれに“SUSIE”の第 100 フレーム以降のフレームを加えている。図 7 から、提案方式が常に従来方式よりも少ない平均二乗誤差で符号化を行っていることがわかる。平均ビットレート 5 Mb/s

図 8 ビットレート配分の時間変化
Characteristics of assigned bitrate for each encoder.

のときで、エンコーダ 3 台の場合、従来方式に対して MSE が約 35% 減少、SNR でみれば約 2 dB 向上している。したがって、エンコーダ間でのビットレート配分がシステム全体としての性能向上につながったことがわかる。一方、エンコーダ 4 台の場合では、MSE が約 45% 減少、SNR で約 2.7 dB の向上となり、エンコーダ 3 台の場合に対してかなり優れた結果となった。この結果から、符号化の易しい画像の数が増えるほどマルチエンコーダシステムの性能が向上することがわかる。さらに、理論値とシミュレーション結果がほぼ一致していることから、前章で行った考察が成り立っていることを示している。

次に、各エンコーダに割り当てられたビットレートの時間変化を図 8 に示す。図 8 では、平均ビットレートを 5 Mb/s としている。それゆえ、各エンコーダの第 1 フレームの割り当てビットレートが 5 Mb/s となっている。図 8 で最も特徴的なのは、画像“POPPLE”に対して割り当てられたビットレートの変化である。“POPPLE”は 75 フレームあたりまでは平坦な背景の中で鳥かごが回っており、それ以降のフレームでは回る鳥かごへズームアップする画像である。したがって、75 フレームまでは比較的符号化の易しい画像であるのに対して、75 フレーム以降は難しいシーンとなる。図 8 から、このシーンの移り変わりを反映して、割り当てられるビットレートが 75 フレームあたりから増加していることがわかる。

このような画像の性質に応じたビットレート配分を行うことで、画質の極度な劣化を防ぐ効果が現れる。それを示したのが図 9 である。図 9 では図 8 と同じ条件でシミュレーションを行い、画像のマクロブロック

図9 符号化歪み特性
Characteristics of coding distortion.

(16 × 16 画素) ごとに SNR を計算し、ある SNR 以下のマクロブロック数が全体の何%存在するかをグラフに示したものである。図9より、提案方式では 35 dB 以下の劣化を従来方式（すべて 5 Mb/s で符号化）に比べて大幅に抑えていることがわかる。

6. む す び

複数の映像プログラムを効率良く一括して伝送するマルチエンコードシステムに関して検討を行った。

まず、比較的少数のエンコード数においても多重化効果を得るために、各エンコードを統一的制御の下に動作させるシステム構成を提案した。次に、映像符号化に広く用いられている MC + DCT 符号化方式において、ビットレート対 MSE の関係がベキ乗の関係で表現されることを示し、各エンコードで発生する MSE の総和を最小にするためのエンコード間ビットレート配分方式を導いた。また、時間的に変化する映像符号化特性に適応させるために、ビットレート配分を更新する方式を提案した。この提案方式を実映像に適用した計算機シミュレーション検証の結果、提案方式はシステム全体の符号化効率を向上させるだけでなく、符号化画像の局所的な画質の極度な劣化も抑制する効果のあることがわかった。

本論文では符号化歪みの尺度として MSE を用いたが、すべての議論は視覚特性など各種の重み付けを行った MSE¹⁰⁾に対しても全く同様に適用できる。

今後の課題は、異なる品質要求を持つ映像プログラムの高効率多重化伝送方式、および本論文では扱わなかった各エンコードからの符号化ビットストリームの多重・分離・ATM 伝送各方式の検討である。

日頃御指導いただく KDD 村上研究所長、古賀シニアグループリーダーに感謝致します。

【参 考 文 献】

- 1) Yamazaki, Wada, Takishima, Wakahara: "ATM Networking and Video-Coding Techniques for QOS Control in B-ISDN", IEEE Trans. CSVT-3, 3 (1993)
- 2) 岸野, 真鍋, 安田: "パケット符号化を考慮した映像信号の統計的性質", Proc. PCSJ'87, 8-5 (1987)
- 3) Koga, et al: "Statistical Performance Analysis of an Inter-frame Encoder for Broadcast Television Signals", IEEE Trans., COM-29, 12 (1981)
- 4) 酒澤, 和田: "マルチエンコードシステムにおける符号化制御に関する検討", テレビ学技報 ICS94-47, 18, 39 (1994)
- 5) 酒澤, 和田: "MPEG-2 マルチエンコードシステムにおけるビットレート配分に関する検討", 1994 テレビ年次大, 11-11 (1994)
- 6) ISO/IEC JTC1/SC29/WG11: MPEG93/457 (Test Model 5)
- 7) Farvardin, Modestino: "Optimum Quantizer Performance for a Class of Non-Gaussian Memoryless Sources", IEEE Trans., IT-30, 3 (1984)
- 8) 酒澤, 滝嶋, 和田: "MC+DCT 符号化器におけるレート歪み特性とそのマルチエンコードシステムへの応用", 信学技報, IE94-127 (1995)
- 9) Jayant, Noll: Digital Coding of Waveforms, Prentice-Hall (1984)
- 10) 浜田, 松本: "画像の局所変化度による雑音マスキング効果を考慮した直交変換係数の最適量子化法", 信学論 B-I, J75-B-I, 12 (1992)

さかざわしげゆき

酒澤茂之 1990 年, 神戸大学工学部電子工学科卒業。1992 年, 同大学院修士課程修了。同年, KDD に入社, 研究所勤務。現在, 同研究所マルチメディア通信グループ主任。動画像符号化方式および ATM や LAN 環境における映像通信システムの研究・開発に従事。正会員。

たかしやすひろ

滝嶋康弘 1986 年, 東京大学工学部電気工学科卒業。1988 年, 同大学院修士課程修了。同年, KDD に入社, 研究所勤務。現在, 同研究所マルチメディア通信グループ主査。情報理論の研究, 動画像符号化の研究・開発・標準化に従事。特に ATM 符号化, 低レート画像符号化技術の応用に従事。正会員。

わだまさひろ

和田正裕 1981 年, 東京大学大学院電気工学専攻修士課程修了。同年, KDD に入社。1983 年より, 同社研究所において低レートビデオ通信方式, テレビ電話・会議システム, ATM ビデオ通信方式, VoD システムの研究開発・標準化に従事。現在, 同研究所マルチメディア通信グループ主任研究員。工学博士。正会員。

はとりよしり

羽鳥好律 1971 年, 東京大学工学部電気工学科卒業。同年, KDD に入社。以来, 画像信号の高効率符号化方式, 通信方式の研究に従事。現在, 同社研究所開発プロジェクトグループグループリーダー。工学博士。正会員。