

階層配分符号量に伴う画質変動を考慮した階層符号化符号量制御方式の検討

笠井 裕之[†] 児玉 明^{††} 富永 英義[†]

[†] 早稲田大学 理工学部 電子・情報通信学科
〒169-8555 東京都新宿区大久保 3-4-1

^{††} 広島大学地域共同研究センター
〒730-0046 広島県東広島市鏡山 3-10-31

kasai@tom.comm.waseda.ac.jp

動画階層符号化方式として、各階層符号量に伴う階層品質の特性解析と階層化特性を考慮した動画階層符号化符号量制御方式の検討を行なう。動画階層符号化方式として MPEG-2 SNR スケーラビリティを採用し、階層配分符号量に伴う画質変動特性をシミュレーション実験から明らかにする。特に符号化器構成法として Double Loop, Single Loop の 2 方式を比較し、その特性を示す。そして両方式の特性から、階層配分符号量の配分率を階層画質特性に基づいて動的に変動・制御することで、階層劣化を極力抑制することが可能な階層符号化符号量制御方式を提案し、その制御法を明らかにする。最後に TM5 方式と比較することにより、提案方式の有効性を示す。

キーワード 階層配分符号量, 符号量制御, SNR スケーラビリティ, TM5

A Study on the Video Scalability Rate Control Method considering Coding Bits Allocation Characteristics

Hiroyuki KASAI[†] Mei KODAMA^{††} Hideyoshi TOMINAGA[†]

[†] Dept. of Electronics, Information and Communication Engineering, WASEDA University
3-4-1 Ohkubo Shinjuku-ku, Tokyo, 169-8555 JAPAN

^{††} Center for Technology Research & Development, Hiroshima University
3-10-31 Kagamiyama, Higashi-Hiroshima, Hiroshima, 739-0046 JAPAN

kasai@tom.comm.waseda.ac.jp

In this paper, we study the analysis of the video quality according to the coding bits allocation to the base-layer and enhancement-layer, and the video scalability coding rate control method considering the relation video image and allocated bits. In particular, we use the MPEG-2 SNR Scalability. First of all, we investigate the efficiency of SNR scalability, which includes Double Loop Method and Single Loop Method, from simulation experiments, and we define the characteristic function. Next we propose the coding rate control method from this characteristic. From the simulation experiment, we show the effitiveness in comparison to TM5 rate control method.

keyword Hierarchical Allocated Bits, Rate Control, SNR Scalability, TM5

1. はじめに

動画階層符号化方式に対する検討は以前から行なわれているが⁽¹⁾, 特に各階層符号量と階層品質特性に関して十分な検討は行なわれておらず⁽²⁾, またそれらを考慮した符号量制御方式に関する検討も十分ではないといえない。そこで、本稿では、MEPG-2⁽³⁾で採用されているSNRスケラビリティを用いて、階層配分符号量に伴う画質変動特性を明らかにするとともに、符号化特性に基づいた階層符号量制御方式を提案し、その有効性を示す。

2. SNR スケラビリティ符号化特性

本節では、シミュレーション実験の結果を考察することにより、階層配分符号量の変化に伴うSNRスケラビリティ符号化特性について明らかにする。特にSNRスケラビリティ符号化器の構成手法として(i)Double Loop, (ii)Single Loopの2手法に分けて考察する。Double Loopは、基本階層、高位階層で独立した予測ループを構成するため、基本階層における予測で用いる参照画像は基本階層局所復号画像を用い、高位階層では高位階層局所復号画像を用いる。よって、予測画像の不一致に起因した画質劣化が生じることはない。一方Single Loopは、基本階層の予測のための参照画像が存在せず、基本階層の予測は高位階層の局所復号画像を用いて行なわれる。このため基本階層のみの復号において符号化側と復号側との予測画像の不一致が生じ、基本階層画質が大きく劣化する可能性がある。

2.1 シミュレーション実験

本節では、MPEG-2 Test Model 5(以下TM5⁽⁴⁾と呼ぶ)に基づき、符号化シミュレーション実験を行なう。テストシーケンスとしてMobile&Calendar, Flower Garden, Table Tennis, 画像フォーマットはITU-R BT.601 4:2:0, 輝度信号は704[pel]×480[line], 色差信号は352[pel]×240[line]を使用した。符号化フレーム数は150フレーム、符号化レートは9[Mbps]であり、基本階層を3.0[Mbps]から8.8[Mbps]まで変動させた。ここで、目標符号化レートに対する目標基本階層目標符号化レートの割合を ρ_{Target} (ただし $0 \leq \rho_{Target} \leq 1.0$)と定義する。

Double Loopにおける ρ_{Target} と高位階層画像SN比の関係を図1に、Single Loopにおける ρ_{Target} と高位階層画像SN比の関係を図2に、両方式の ρ_{Target} と基本階層画像SN比の関係を図3に示す

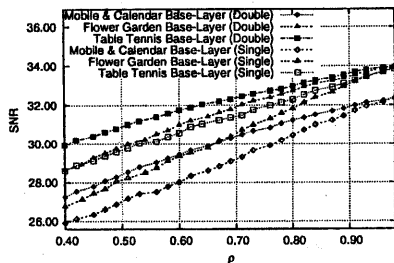


図3: 符号量配分率と基本階層SN比の関係

高位階層特性に関しては、図1, 2に示すように、Double Loop, Single LoopともにIピクチャでは、符号量配分率の増加に伴い高位階層SN比は減少し、配分率が約60%~70%で極小値を示す。その後配分率の増加に伴い高位階層

SN比は増加する。つまり中間的な符号量配分を行なう場合には、比較的大きな階層化劣化を引き起こしてしまい、一方でどちらかの階層に極端な符号量配分をした場合、階層化劣化を抑制することが可能であることが分かる。これは、極端な符号量配分をした場合、階層符号化にも関わらず仮想的に非階層符号化に近くなり、階層化劣化が減少するためであると考えられる。また、P,Bピクチャに関しても同様の傾向を示し、約60%付近で最小値を示すことが分かる。しかしながら、Double LoopとSingle Loopの結果と比較した場合、高配分率では同等のSN比を示すが、中間配分率から低配分率においてはDouble Loopは明らかに劣化していることが分かる。これからSingle Loopでは参照画像を高位階層画像を用いるため、基本階層の品質に左右されず、低配分率においても良好な品質を提供できることが分かる。

一方、基本階層に関しては図3に示すように、Double Loop, Single Loopともに各シーケンスとも配分率の増加とともに一様に増加しており、これは割り当て符号量の増加と品質が比例することを示している。しかしながら、Double LoopとSingle Loopの結果と比較した場合、高配分率では同等のSN比を示すが、中間配分率から低配分率においては明らかにDouble Loopが優位な結果を示す。本傾向は、Double Loopの基本階層における参照画像の不一致に起因した画質劣化が生じていないことを示している。

2.2 階層化劣化関数D(rho)と特性関数phi(rho)の導出

階層化劣化関数D(rho)を式(1)に定義し、2.1のシミュレーション結果からD(rho)を図4に示す。

$$D(\rho) = 10 \left(\frac{SNR(\rho) - SNR_{非階層}}{10} \right) \quad (1)$$

図4からD(rho)の近似曲線として特性関数phi(rho)を定義する。図4には2次関数により近似した曲線を示した。ここで、特性関数phi(rho)の極小値を示すrhoをrho_0と定義する。

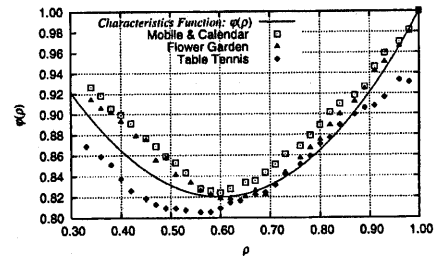


図4: 階層化劣化関数と特性関数

2.3 考察

2.から、符号化特性についてまとめる。

- (a) Double Loop, Single Loopともに約60%付近で階層化劣化を引き起こす。
- (b) Double Loop: 低・中間配分率において,
 - (i) 基本階層: ドリフトが生じないため高品質
 - (ii) 高位階層: 基本階層品質低下に伴う品質劣化
- (c) Single Loop: 低・中間配分率において,
 - (i) 基本階層: ドリフトによる品質劣化
 - (ii) 高位階層: 予測における高位階層局所復号画像の使用による品質劣化の抑制, 高品質画像の提供

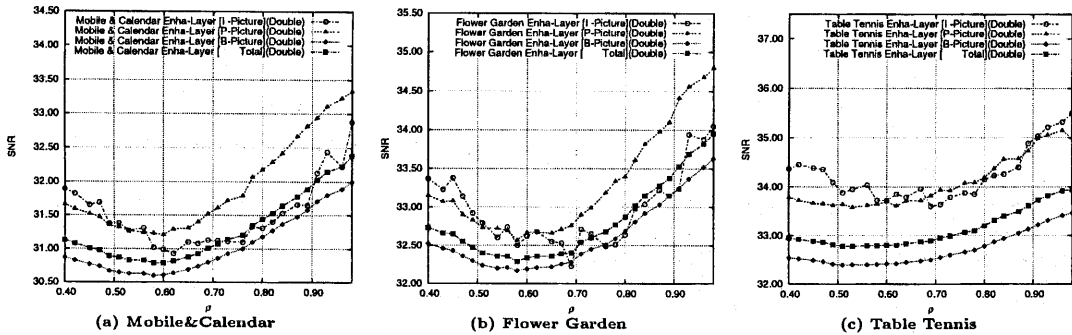


図 1: 符号量配分率と高位階層 SN 比の関係 (Double Loop)

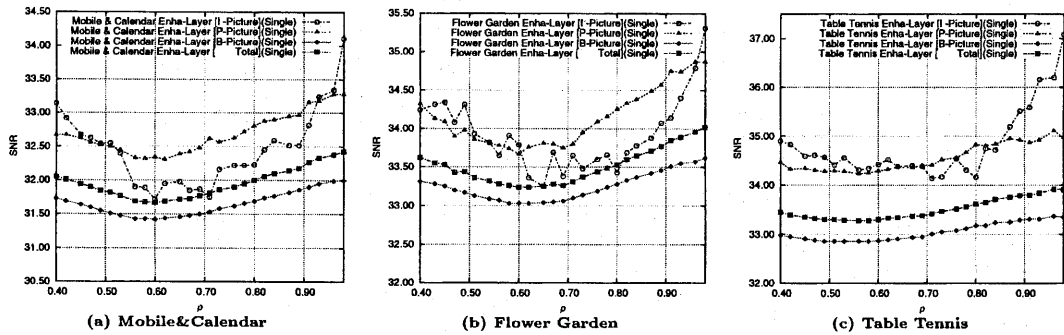


図 2: 符号量配分率と高位階層 SN 比の関係 (Single Loop)

3. 配分符号量の変動を考慮した符号量制御方式の提案

2.の考察から、配分符号量の変動に伴う階層品質特性を考慮した符号量制御方式を提案し、その概念を述べる。

3.1 提案符号量制御方式の概念

2.1において、各ピクチャにおける符号量配分率により品質が大きく変動することを示し、2.2において、その品質変動は特性関数 $\varphi(\rho)$ に基づくことを示した。そこで、特性関数 $\varphi(\rho)$ に基づいて、各ピクチャにおいて符号量配分率 ρ を動的に変動させることにより階層化劣化を抑え、性能改善を実現する階層符号化符号量制御方式を提案する。本提案方式は、以下の基本的方針に基づくものとする。

- 方針 (a) GOP 単位での目標符号量配分率 ρ_{Target} を一定に保つ。
- 方針 (b) 極端な符号量配分を実現することにより階層化劣化を抑える。
- 方針 (c) 基本階層でのドリフトを極力抑える。
- 方針 (d) 符号量配分率の更新により各ピクチャの劣化が生じない。

以上の方針に基づいて、以下の配分率制御を実現する。また、この様子を図5に示す。

制御 1 GOP, ピクチャ単位の符号量制御は TM5 の Step 1 に基づき、各ピクチャ単位での符号量制御では、基本階層、高位階層の符号量の総符号量は TM5 制御に基づいて行なう [方針 (a)]。

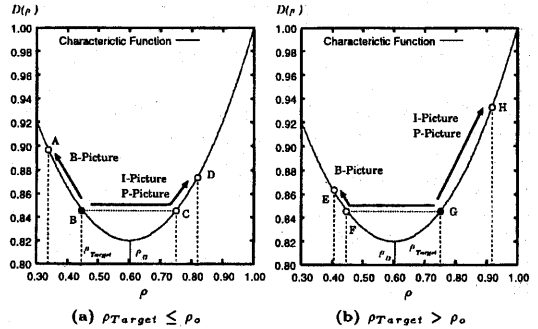


図 5: 提案符号化制御方式概念図

制御 2 I,P ピクチャでは、符号量配分率 ρ の増える方向、つまり 1.0 に近い値になるように更新する。一方、B ピクチャでは、符号量配分率 ρ の減少する方向、つまり 0.1 に近い値になるように更新する。これにより各ピクチャでの階層化劣化を極力抑えると同時に、基本階層における参照画像品質が向上するのでドリフトを抑えることが可能となる [方針 (b),(c)]。

制御 3 $\rho_{Target} \leq \rho_0$ の場合における I,P ピクチャの配分率の更新、及び $\rho_{Target} > \rho_0$ の場合における B ピクチャの配分率の更新にあたっては、以下の式 (2) を満たす ρ へ変換することにより品質劣化の発生を防ぐことが可能となる。但し配分率の更新を行なうピクチャのタイプに関しては、消費符号量の割合から、P ピクチャの

更新を積極的におこない、Pピクチャのみで実現不可能な場合に限ってはIピクチャの更新を行なうものとする。[方針(d)].

$$\begin{cases} \rho \geq \varphi^{-1}\{\varphi(\rho)\} & \text{if } \rho_{Target} \leq \rho_0 \\ \rho \leq \varphi^{-1}\{\varphi(\rho)\} & \text{if } \rho_{Target} > \rho_0 \end{cases} \quad (2)$$

4. 提案符号量制御方式

本節では、具体的に符号量制御方式を述べる。まず MPEG-2 Test Model 5(以下 TM5⁽⁴⁾と呼ぶ)の符号量制御方式を概説する。次に、符号量制御を行なう前処理として最適な符号量配分率 ρ_{ideal} の算出手法と、 ρ_{ideal} を用いた基本階層/高位階層目標符号量 $T_{(i,p,b)}(l,h)$ の導出手法を示す。ここで (l,h) は基本階層、高位階層を示す添字である。

4.1 TM5 符号量制御方式

TM5における符号量制御アルゴリズムは3ステップから構成されており、Step 1では、以下の式(3)に従って各ピクチャに対する目標符号量 $T_{(i,p,b)}$ を決定する。

$$\begin{cases} T_i = R / (N_i + \alpha + \beta) \\ T_p = R / (N_p + \gamma) \\ T_b = R / (N_b + \delta) \end{cases} \quad (3)$$

式(3)において、 R はGOP内の未符号化ピクチャに対して割り当てられる符号量、 N_p, N_b はGOP内の未符号化P,Bピクチャの枚数である。但し $(\alpha, \beta, \gamma, \delta)$ は未符号化ピクチャの当該ピクチャの相当枚数を示すパラメータであり、例えば α は、 N_p 枚のGOP内の未符号化PピクチャがIピクチャ何枚分の符号量に相当するかを示すものである。

4.2 最適符号量配分率 ρ_{ideal} と目標符号量 $T_{(i,p,b)}(l,h)$ の導出

本符号量制御方式は、最終的に実現すべき目標符号量配分率 ρ_{Target} を目標として、各ピクチャ符号化前に、特性関数 $\varphi(\rho)$ 及び式(3)から、ピクチャ単位の最適符号量配分率 ρ_{ideal} を算出し、算出した ρ_{ideal} を用いて最適な符号量配分制御を実現する。但し、既に符号化したP,Bピクチャの枚数を N'_p, N'_b とし、それぞれ $N'_p = N/M - 1 - N_p, N'_b = N - N/M - N_b$ により算出する。また、Bピクチャにおいて基本階層で必要とされる最低限の符号量を $T_{min}(l)$ 、I,Pピクチャにおいて高位階層で必要とされる最低限の符号量を $T_{min}(h)$ と定義する。そして実際に符号化に要した実符号量を $R_{(i,p,b)Real}(l,h)$ と定義する。さらに $T_{(i,p,b)}$ は式(4)を満たすものとする。

$$T_{(i,p,b)} = T_{(i,p,b)}(l) + T_{(i,p,b)}(h) \quad (4)$$

4.2.1 $\rho_{Target} \leq \rho_0$ の場合

■ Iピクチャ

符号量配分率の更新は行なわず最適符号量配分率 ρ_{ideal} を式(5)に定義する。

$$\rho_{ideal} = \rho_{Target} \quad (5)$$

式(5)から目標符号量は式(6)により導出する。

$$\begin{cases} T_i(l) = T_i \times \rho_{ideal} \\ T_i(h) = T_i \times (1 - \rho_{ideal}) \end{cases} \quad (6)$$

■ Pピクチャ

Pピクチャでは、基本階層において、既にI,Pピクチャで過度に消費した符号量 R_{past}^+ 、Bピクチャにおいて節約した符号量 R_{past}^- 、及び未符号化Bピクチャにおいて節約されることが予測される符号量 R_{pre}^- を考慮して、符号量配分率の更新を行なう。

- 既にI,Pピクチャにより過度に消費した符号量: R_{past}^+

$$R_{past}^+ = \{T_{iReal}(l) - T_i \times \rho_{Target}\} \times N'_p + \sum \{T_{pReal}(l) - T_p \times \rho_{Target}\} \quad (7)$$

- 既にBピクチャにおいて節約された符号量: R_{past}^-

$$R_{past}^- = \sum \{T_b \times \rho_{Target} - T_{bReal}(l)\} \quad (8)$$

- 未符号化Bピクチャにおける節約可能符号量: R_{pre}^-

$$R_{pre}^- = T_p \times \gamma \times \rho_{Target} - T_{min}(l) \times N_b \quad (9)$$

式(7)~(9)を用いて、式(10)を満たす ρ を最適符号量配分率 ρ_{ideal} と定義する。

$$T_p \times \{\rho - \rho_{Target}\} = (R_{pre}^- + R_{past}^- - R_{past}^+) / N_p \quad (10)$$

式(10)から目標符号量は式(11)により導出する。

$$\begin{cases} T_p(l) = T_p \times \rho_{ideal} \\ T_p(h) = T_p \times (1 - \rho_{ideal}) \end{cases} \quad (11)$$

■ Bピクチャ

目標符号量を式(12)により導出する。

$$\begin{cases} T_b(l) = T_{min}(l) \\ T_b(h) = T_b - T_{min}(l) \end{cases} \quad (12)$$

4.2.2 $\rho_{Target} > \rho_0$ の場合

■ Iピクチャ

- 未符号化Pピクチャによる消費予測符号量: R_{pre}^+

$$R_{pre}^+ = T_i \times \alpha - N_p \times T_{min}(h) - T_i \times \alpha \times \rho_{Target} \quad (13)$$

- 未符号化Bピクチャによる節約予測符号量: R_{pre}^-

$$R_{pre}^- = T_i \times \beta \times \rho_{Target} - T_{min}(l) \times N_b \quad (14)$$

上記2式からIピクチャ使用ON/OFFの判定を行なう。

(a) $R_{pre}^+ \geq R_{pre}^-$ の場合

符号量配分率の更新を行ない、Iピクチャにおける目標符号量を式(15)により導出する。

$$\begin{cases} T_i(l) = T_i - T_{min}(h) \\ T_i(h) = T_{min}(h) \end{cases} \quad (15)$$

ここで最適符号量配分率は式(16)に定義される。

$$\rho_{ideal} = \{T_i - T_{min}(h)\} / T_i \quad (16)$$

(b) $R_{pre}^+ < R_{pre}^-$ の場合

符号量配分率の更新は行なわず $\rho_{ideal} = \rho_{Target}$ であり、I目標符号量は式(6)と同式により算出する。

■ Pピクチャ

Pピクチャの目標符号量を式(17)により導出する。

$$\begin{cases} T_p(l) = T_p - T_{min}(h) \\ T_p(h) = T_{min}(h) \end{cases} \quad (17)$$

ここで最適符号量配分率は式(18)に定義する。

$$\rho_{ideal} = \{T_p - T_{min}(h) \times N_p\} / T_p \quad (18)$$

■ Bピクチャ

- 既にI,Pピクチャにより過度に消費した符号量 R_{past}^+ は式(7)と同式。
- 既にBピクチャにおいて節約された符号量 R_{past}^- は式(8)と同式。
- 未符号化Pピクチャにおける消費予測符号量: R_{pre}^+

$$R_{pre}^+ = T_b \times \delta - T_{min}(h) \times N_p - T_b \times \delta \times \rho_{Target} \quad (19)$$

式(7),(8),(21)を用いて、式(20)を満たす ρ を最適符号量配分率 ρ_{ideal} と定義する。

$$T_b \times \{\rho - \rho_{Target}\} = (R_{past}^+ - R_{past}^- + R_{pre}^+) / N_b \quad (20)$$

式4.2.2から目標符号量は式(21)により導出する。

$$\begin{cases} T_b(l) = T_b \times \rho_{ideal} \\ T_b(h) = T_b \times (1 - \rho_{ideal}) \end{cases} \quad (21)$$

5. 提案方式の評価

提案方式の性能評価及び有効範囲に対する評価を行う。シミュレーション評価実験条件は2.と同様である。

5.1 符号量配分率の変動

提案符号量制御方式における符号量配分率 ρ の変動を図6に示す。図6に示すようにTM5方式と比較して符号量配分率が激しく変動しており、I,Pピクチャでは高配分率、Bピクチャでは低配分率を示している。 ρ_{Target} の相違による傾向として、 $\rho_{Target} < \rho_0$ の場合には、全てのピクチャにおいて ρ_{ideal} が低く設定されており、一方、 $\rho_{Target} > \rho_0$ の場合には ρ_{ideal} が高く設定されていることが分かる。特に $\rho_{Target} < \rho_0$ におけるBピクチャの ρ が最も低い値を示していること、 $\rho_{Target} > \rho_0$ におけるI,Pピクチャの ρ が最も高い値を示していることから、本提案方式の制御が $\rho_{Target} < \rho_0$ においてBピクチャの $T_b(l)$ を $T_{min}(l)$ に、 $\rho_{Target} > \rho_0$ においてI,Pピクチャの $T_{i(p)}b(h)$ を $T_{min}(h)$ に設定していることが理解できる。

5.2 SN比の変動

各符号量配分率における高位階層のSN比を図7に示す。また2.において最大劣化を示した配分率 $\rho = 0.60$ におけるSN比を表1に示す。また、各フレームにおける高位階層のSN比の変動を図8に示す。

表1: 各方式におけるSN比の関係 ($\rho_{Target} = 0.60$)

方式	Double	Single	提案方式	MP@ML
MOBL	30.79	31.67	32.41	32.51
FLOW	32.34	33.24	33.87	34.11
TABL	32.80	33.30	33.80	34.22

図7から、提案方式の効果が現れていることが分かり、Double Loop及びSingle Loopの階層化劣化が抑えられていることが理解できる。また表1においてもTM5方式に比べ、階層化劣化を大きく抑えていることが分かる。さらに図8から、各フレームにおいてTM5方式を上回っており、両シーケンスともに1番目のGOPではそれほど大きな効果はないものの、それ以降のGOPにおいては大きく改善されていることが分かる。

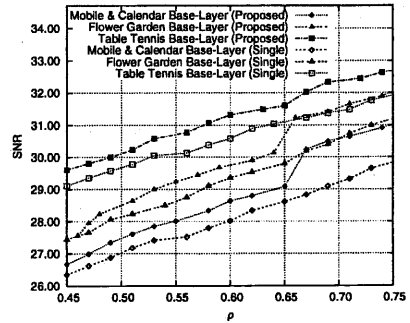


図9: 符号量配分率と基本階層SN比の関係

一方、基本階層の結果を図8及び図9に示す。図9から、全てのシーケンスにおいて品質が向上しており、動き補償用参照画像として用いられるPピクチャの基本階層へ多くの符号量を配分した効果が現れていることが分かる。特にIピクチャの配分符号量の更新を行なった $\rho = 0.65 \sim 0.70$ 近辺においては著しく品質が向上しており、Iピクチャの基本階層へ多くの符号量を配分した効果が現れていることが同様理解できる。また図8から、BピクチャにおいてTM5 Single Loop方式より劣化しているフレームが見うけられるが、I,Pにおいて大きく改善されており、全体として品質が向上していることが分かる。

5.3 提案方式の有効符号量配分率の範囲

本節では、提案方式が有効な符号量配分率 ρ の範囲を理論的に導出する。ここで $\rho_{Target} = x$, $\varphi^{-1}\{\varphi(\rho_{Target})\} = y$ と定義し、各場合における ρ_{Target} の下限値 $\rho_{TargetMIN}$ 、上限値 $\rho_{TargetMAX}$ を導出する。

5.3.1 有効下限値

I,Pピクチャの ρ を更新する場合においては、Bピクチャへの配分符号量が $T_{min}(l)$ であり、I,Pピクチャの配分率 $\rho(D$ 点)がC点と一致する場合である。この時、以下の式(22)を満たす x が $\rho_{TargetMIN}$ に相当する。

$$T_i(y - x) + T_i\alpha(y - x) = T_i\beta x - T_{min}N_b \quad (22)$$

また、Pピクチャの ρ のみを更新する場合(Iピクチャ:B点、Pピクチャ:C点)は、式(22)の左辺第1項が無い場合、Iピクチャの ρ のみを更新する場合(Iピクチャ:C点、Pピクチャ:B点)は式(22)の左辺第2項が無い場合である。

5.3.2 有効上限値

I,Pピクチャの ρ を更新する場合においては、I,Pピクチャへの配分符号量が $T_{min}(h)$ であり、Bピクチャの配分率 $\rho(E$ 点)がF点と一致する場合である。この時、以下の式(23)を満たす x が $\rho_{TargetMAX}$ に相当する。

$$\{T_i - T_{min}(h) - T_i x\} + \{(T_i\alpha - T_{min}(h)N_p) - T_i\alpha x\} = T_i\beta x - T_i\beta y \quad (23)$$

5.3.3 有効範囲理論値

式(22)、(23)に基づいて $\rho_{TargetMIN}$ 、 $\rho_{TargetMAX}$ を導出した。表2に算出値を示す。

表2から、 ρ を更新するピクチャタイプを適宜選択することにより、 $0.36 \leq \rho_{Target} \leq 0.76$ の範囲において、本符号量制御方式が有効であることが分かる。

□ 階層配分符号量に伴う画質変動を考慮した階層符号化符号量制御方式の検討

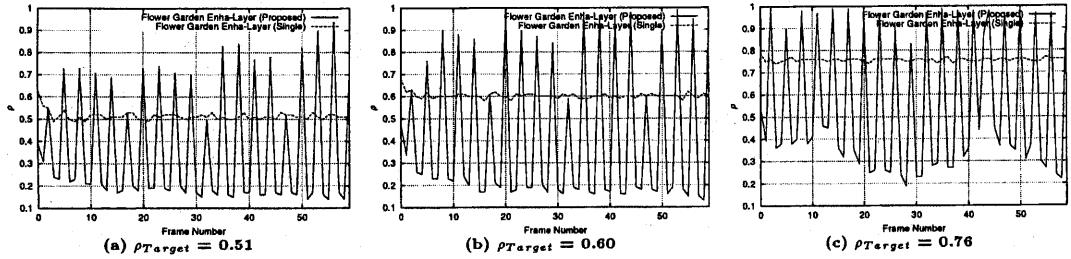


図 6: 符号量配分率の変動

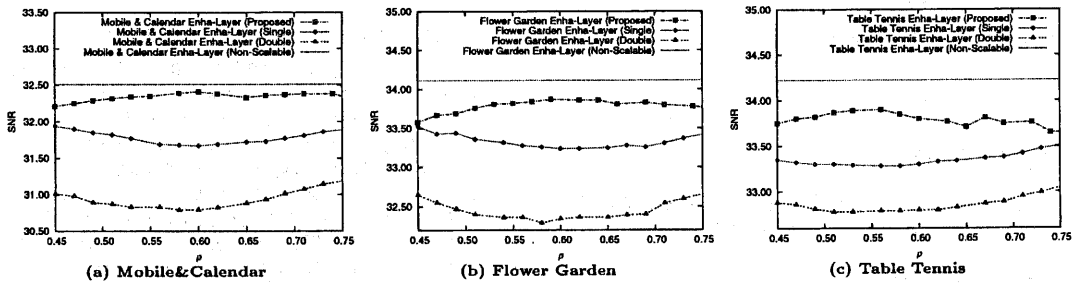


図 7: 符号量配分率と高位階層 SN 比の関係

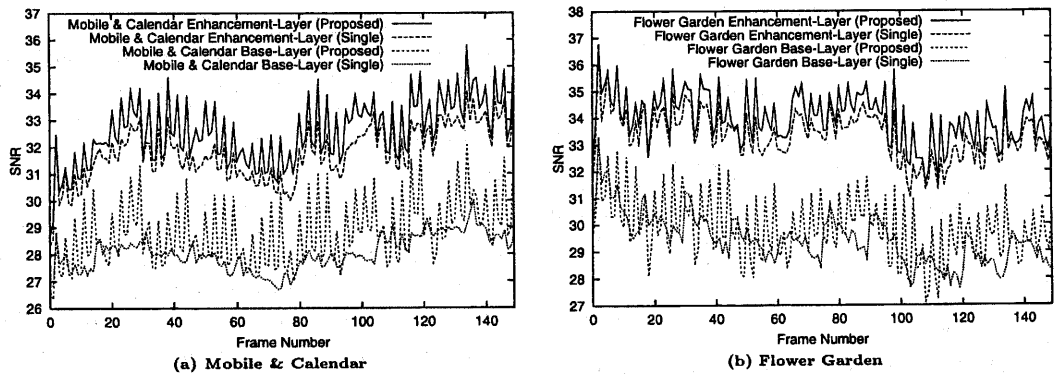


図 8: 各フレームにおける高位階層 / 基本階層 SN 比の変動

表 2: $\rho_{TargetMIN}$ と $\rho_{TargetMAX}$ 理論値

Case	$\rho_{TargetMIN}$	$\rho_{TargetMAX}$
I, P の ρ を更新	0.47	0.76
P の ρ を更新	0.41	0.71
I の ρ を更新	0.36	0.68

6. まとめ

本稿では、動画階層符号化として MPEG-2 SNR スケラビリティを採用し、階層配分符号量と基本 / 高位階層品質の特性について検討し、高位階層では Single Loop, 基本階層では Double Loop が有効であることを示した。そして階層配分符号量による画質変動を考慮して、符号量配分率をピクチャ単位で動的に制御することで、階層特性に基づいた動画階層符号化符号量制御方式を提案した。その結果、提案方式が、TM5 方式の性能改善を実現し、階層化に伴って発生する階層化劣化を抑制する方式であることを示し、そ

の有効性を示した。今後は、配分符号量の更新を行なうピクチャタイプによる品質の変動、及びそれらを適応的に変更するための判定法について検討する。

参考文献

- (1) 杉原 他: “動画階層符号化におけるスケラビリティに関する研究”, PCSJ 92, pp. 165-168 (1992).
- (2) H. KASAI, M. KODAMA and H. TOMINAGA: “A Consideration on the Efficiency of Image Scalability from Coding Bits Allocation”, PCS'97, pp. 441-444 (1997).
- (3) ISO-IEC 13818-2, IS: “Generic Coding of Moving Picture and Associated Audio Recommendation H.262” (1995).
- (4) ISO-IEC JTC1, MPEG 93/N225b: “Test Model 5 /draft” (1993).