

ビットストリーム変換による 階層間独立スケーラブル映像符号化方式の検討

涌井 道子[†] 永吉 功^{††,†††} 花村 剛^{†††,††} 富永 英義^{††,†††}

[†] 早稲田大学大学院 理工学研究科 〒169-8555 東京都新宿区大久保 3-4-1

^{††} 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 メディアグルー 〒169-0072 東京都新宿区大久保 2-4-12 新宿ラムダックスビル 8 階

^{††††} 早稲田大学 国際情報通信研究センター 〒169-0051 東京都新宿区西早稲田 1-3-10

E-mail: †{miti,isao,tominaga}@tom.comm.waseda.ac.jp, ††hana@mediaglu.co.jp

あらまし 本稿では、階層間の依存関係を排除したスケーラブル映像符号化方式を検討する。まず、従来のスケーラブル映像符号化方式を用いた階層型映像伝送における課題を整理し、符号化方式に求められる基本要件を示す。次に、上記課題を解決するため、階層間の依存関係を排除し、任意の階層の組合せによる映像再生を実現する符号化処理アルゴリズムを提案し、その詳細を説明する。最後に評価実験より、依存関係を排除した階層構造においても、符号量肥大化の回避が実現可能であることを示し、提案方式の有効性を示した。

キーワード 階層間独立性、スケーラブル映像符号化、ビットストリーム変換、MPEG

A Study on a Scalable Video Coding Scheme with Independent Layers by Transcoding

Michiko WAKUI[†], Isao NAGAYOSHI^{††,†††}, Tsuyoshi HANAMURA^{†††,††}, and Hideyoshi
TOMINAGA^{††,†††}

[†] Graduate School of Sci. and Eng., WASEDA University, 3-4-1 Ohkubo, Shinjuku-ku, Tokyo,
169-8555 Japan

^{††} GITS, WASEDA University, 29-7 building 1-3-10 Nishi-Waseda, Shinjuku-ku, Tokyo, 169-0051 Japan

^{†††} Media Glue Corp., Ramdax Bldg. 8th floor, 2-4-12 Ohkubo, Shinjuku-ku, Tokyo, 169-0072 Japan

^{††††} GITI Waseda University, 1-3-10 Nishi-Waseda, Shinjuku-ku, Tokyo, 169-0051 Japan

E-mail: †{miti,isao,tominaga}@tom.comm.waseda.ac.jp, ††hana@mediaglu.co.jp

Abstract In this paper, we describe a study on a scalable video coding scheme with independent layers through transcoding. First, we discuss about scalable video transmission by traditional scalable video coding scheme. Second, we explain the proposed method. It converts non-layered bitstream to independent layer bitstream by transcoding. At last, from the simulation experiments, we show that the proposed coding algorithm can prevent bit-count to become large, and show the effectiveness of proposed coding scheme.

Key words Layer Independency, Scalable Video Coding, Bitstream Transcoding, MPEG

1. ま え が き

デジタル映像を取り巻くさまざまなサービスに対する要求が高まる中、同一映像情報から利用環境に応じてさまざまな品質を選択的に提供可能な技術として、筆者らは階層型映像伝送方式を検討している。本方式は、ビットストリーム変換により非階層映像情報から基本映像品質を与える基本階層と基本品質

を向上させる拡張階層を生成し伝送することで、スケーラビリティを実現する [1] [2]。一方、階層映像情報の伝送効率に着目すると、階層構造生成にかかる符号化器・復号器の処理量の増大、ネットワークルータ等における伝送優先制御の複雑性が課題であることが明らかになっている [3]。

そこで本稿では、符号化方式の観点で階層型映像伝送における課題を指摘し、スケーラブル映像符号化方式に対する基本要

件を整理する．また，拡張階層間での依存関係を排除することにより，上記課題を解決する新たなスケーラブル映像符号化方式を提案し，評価実験により有効性を示す．

2. 階層型映像伝送における課題

階層型映像伝送における課題を以下に示す [3]．

- ・ 階層間の依存関係により下位階層から順に優先制御の必要性
- ・ 階層数の増加に伴うサーバ，ネットワーク負荷の増大
- ・ 中間階層損失時において損失階層以上の階層符号の復号が不可能

上記課題は，伝送する映像情報構造に起因すると考えられる．階層型映像伝送における階層映像情報を生成する手段として，MPEG-2 では SNR スケーラブル符号化方式 [4] が標準化され，MPEG-4 では画像品質の選択に柔軟性を実現した FGS (Fine Granularity Scalability) [5] が標準化された．また，筆者らは，ビットストリーム変換により階層構造を生成するスケーラブル映像符号化 [1] [2] をこれまで検討してきた．しかし，いずれの方式においても階層間に依存関係が存在し，図 1 のように中間階層を損失した場合，受信符号量に相当する映像品質は提供されない．このような中間階層損失時における耐性の低減を回避するため複雑な伝送優先制御が必要となる．

そこで，上記課題を解決するために階層間の依存関係を排除した映像情報構造を生成することにより高いスケーラビリティを実現する新たな符号化方式の検討が必要となる．本方式を階層間独立スケーラブル映像符号化方式と定義し基本要件を以下にまとめる．

1. 任意の階層の組合せで映像再生可能
2. 符号化効率の向上

3. 階層間独立スケーラブル映像符号化方式

本節では，上記基本要件を満たす具体的な符号化処理方式を述べる．本方式は，すでに符号化済みの非階層映像情報を，階層間の依存関係を排除した階層映像情報に変換する方式であり，

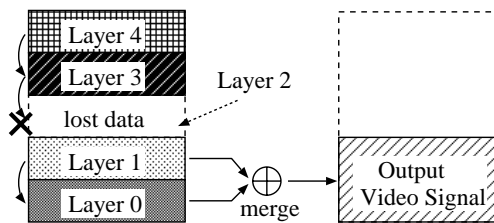


図 1 従来のスケーラブル映像符号化方式

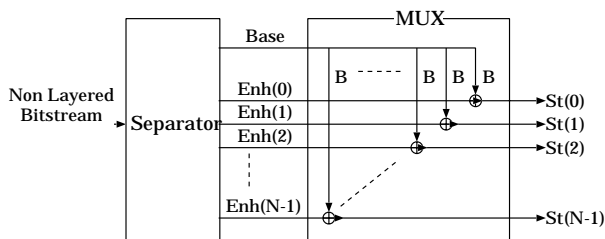


図 2 階層ストリームの生成処理器

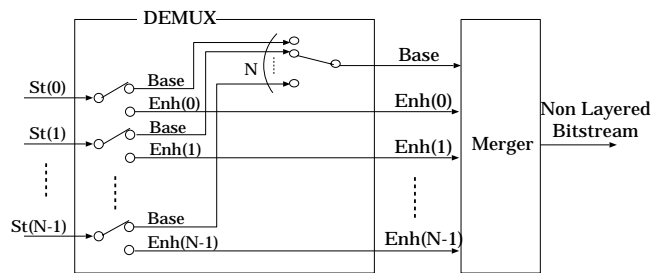


図 3 非階層ストリームの復元処理器

DCT 係数領域トランスコーダ [6] を拡張した方式 [1] において，拡張階層間の依存関係を排除することにより実現する．

3.1 基本原理

単独階層での映像再生を実現し，任意の階層の組合せで映像再生を可能とするため，各階層は，映像再生に最低限必要とされる基本情報と映像の詳細を表現する情報（部分拡張情報）から構成される．基本情報はすべての階層が共通にもち，部分拡張情報はそれぞれ互いに独立した情報である．

3.2 処理器の構成

階層映像情報の生成処理器の構成を図 2 に，複数の階層映像情報から非階層映像情報を復元する処理器の構成を図 3 に示す．また，基本情報，部分拡張情報から構成されるストリームをそれぞれ，基本ストリーム $Base$ ，部分拡張ストリーム $Enh(n)$ と定義する．また，最終的に出力される互いに独立な構造を有するストリームを階層ストリーム $St(n)$ とする．

(1) 階層ストリームの生成

分離器 (Separator) において非階層ストリームを $Base$ と N 本の $Enh(n)$ へ分離する．次に多重化器 (MUX) において $Base$ を複製し，それぞれを $Enh(n)$ と多重化 [7] して N 本の $St(n)$ を出力する．ここで， N を階層ストリーム数と定義する．

(2) 非階層ストリームの復元

分割器 (DEMUX) において $St(n)$ をそれぞれ， $Base$ と $Enh(n)$ に分割し，合成器 (Merger) において $Base$ と N 本の $Enh(n)$ を合成し非階層ストリームを出力する．

3.3 符号化処理方式

本節では，分離器 (Separator)，合成器 (Merger) における $Base$ ， $Enh(n)$ の具体的な生成処理方法について述べる．

3.3.1 非階層構造から階層構造への変換

分離器 (Separator) の構成を図 4 に示す．また，係数上における階層構造への変換方法を図 5 に示す．入力された非階層ビットストリームはマクロブロック (以後，MB) 層以上のデータと係数情報とに分離される．係数情報は可変長復号 (VLD)，

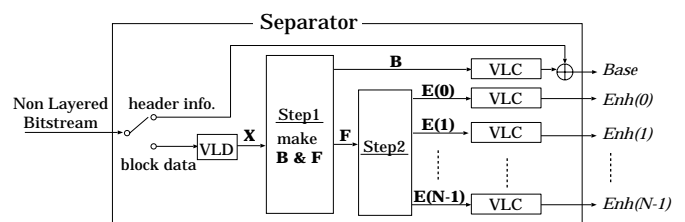


図 4 分離器の構成

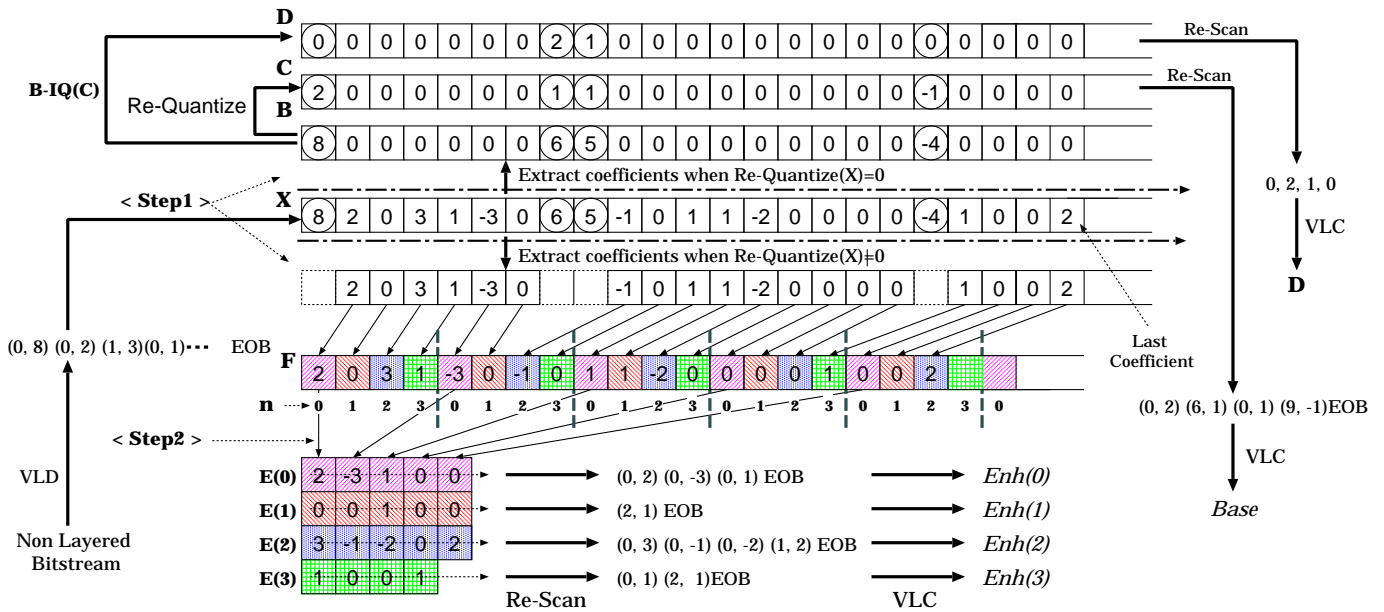


図 5 独立性を有する階層構造への変換方法

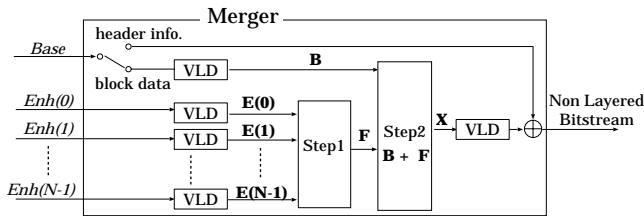


図 6 合成器の構成

ブロック復元処理をした後、1 次元化され量子化係数列 X を生成する。 X は最低限の品質を提供する基本信号 B とそれ以外の情報から構成される拡張信号 F に分離される [1](Step1)。さらに F を N 個の $E(n)$ に分離する (Step2)。 B 、 $E(n)$ はそれぞれ独立に可変長符号化 (VLC) され $Base$ 、 $Enh(n)$ となる。ただし、入力ビットストリームを構成するヘッダー情報は、 $Base$ を構成するヘッダー情報として再利用される。

Step 1 : 基本信号と拡張信号の分離

X において再量子化処理した場合に非零係数となる信号を B 、零係数となる信号を F とする。 B は再量子化され C に変換された後、可変長符号化 (VLC) され $Base$ となる。ただし、再量子化の際に発生する誤差信号 D は可変長符号化 (VLC) され MB 毎に異なる $Enh(n)$ に分配される。

Step 2 : 拡張信号の分離

Step1 において生成された F を N 個の $E(n)$ へ分離する。 F を式 (1) と定義する。

$$F = \{f_0, f_1, f_2, \dots, f_j, \dots\} \quad (0 \leq i < 64) \quad (1)$$

ここで、 j はブロック内の参照順序をジグザグスキャンの順番とした際の係数位置インデックス番号を表す。 F を N 個に区切った区間内の n 番目の係数を部分拡張ストリーム番号が n となる $E(n)$ へ出力する。分離は式 (2) により行なわれる。こ

ここで、 n を階層ストリーム番号と定義する。

$$\begin{cases} E(0) = \{f_0, f_N, f_{2N}, \dots, f_{kN} \dots\} \\ E(1) = \{f_1, f_{1+N}, f_{1+2N}, \dots, f_{1+kN} \dots\} \\ \vdots \\ E(n) = \{f_n, f_{n+N}, f_{n+2N}, \dots, f_{n+kN}, \dots\} \\ \vdots \\ E(N-1) = \{f_{N-1}, f_{N-1+N}, \dots, f_{N-1+kN} \dots\} \end{cases} \quad (2)$$

$E(n)$ 内の k 番目の係数を e_k^n とすると、式 (3) となる。

$$e_k^n = f_{kN+n} \quad (3)$$

N を用いて n, k と j は一意に関係づけられ式 (4) となる。

$$\begin{cases} n = j \bmod N \\ k = \lfloor \frac{j}{N} \rfloor \end{cases} \quad (4)$$

ただし、 $[a]$ は a を越えない最大の整数を、 $a \bmod p$ は a を p で除算した余りを意味する。 $E(n)$ は独立に再スキャンし可変長符号化 (VLC) される。

3.3.2 階層構造から非階層構造への変換

合成器 (Merger) の構成を図 6 に示す。また、係数上における非階層構造への変換方法を図 7 に示す。入力された $Base$ と N 本の $Enh(n)$ をそれぞれ独立に可変長復号 (VLD) し、取得したランレングス係数情報からブロック復元処理を経て 1 次元量子化係数列 B と $E(n)$ に変換する。まず、 N 本の $E(n)$ を合成し F を生成する (Step1)。次に B と F を合成し X を復元する (Step2)。 X は可変長符号化 (VLC) され非階層ストリームとして出力される。ただし、 $Base$ を構成するヘッダー情報は、非階層ストリームを構成するヘッダー情報として再利用される。

Step 1 : 部分拡張信号の合成

$Enh(n)$ を可変長復号 (VLD) し $E(n)$ と D を取得する。次

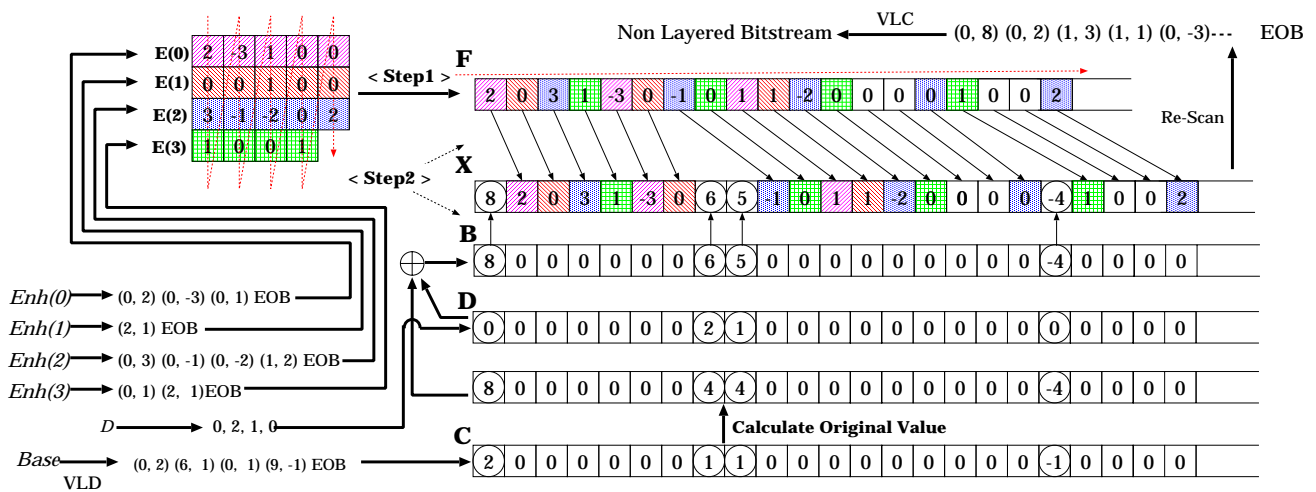


図 7 非階層構造への復元方法

に、階層ストリーム番号 n に基づき $E(n)$ の合成を行ない F を生成する。 $E(n)$ を式 (5) のように定義する。

$$\begin{cases} E(0) = \{e_0^0, e_1^0, e_2^0, \dots\} \\ E(1) = \{e_0^1, e_1^1, e_2^1, \dots\} \\ \vdots \\ E(n) = \{e_0^n, e_1^n, e_2^n, \dots\} \\ \vdots \\ E(N-1) = \{e_0^{N-1}, e_1^{N-1}, e_2^{N-1}, \dots\} \end{cases} \quad (5)$$

$E(n)$ の系列を F と同系列に拡張した量子化係数列 $G(n)$ を式 (6) のように生成する。

$$\begin{cases} G(0) = \{\underbrace{e_0^0, 0, 0, \dots, e_1^0, 0, 0, \dots}_{N}, \dots\} \\ G(1) = \{\underbrace{0, e_0^1, 0, \dots, 0, e_1^1, 0, \dots}_{N}, \dots\} \\ \vdots \\ G(n) = \{\underbrace{0, \dots, e_0^n, 0, \dots, 0, \dots, e_1^n, 0, \dots}_{N}, \dots\} \\ \vdots \\ G(N-1) = \{\underbrace{0, \dots, 0, e_0^{N-1}, 0, \dots, 0, e_1^{N-1}, \dots}_{N}, \dots\} \end{cases} \quad (6)$$

$G(n)$ の j 番目の要素を g_j^n と表すと、 g_j^n は式 (7) のように表される。

$$g_j^n = \begin{cases} e_k^n & \text{for } j = kN + n \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

$G(n)$ を用いて N 個の $E(n)$ を合成した F は式 (8) から生成される。

$$F = \sum_{n=0}^{N-1} G(n) \quad (8)$$

Step 2: 基本信号と拡張信号の合成

Base を可変長復号 (VLD) し取得した C と D から、 B を

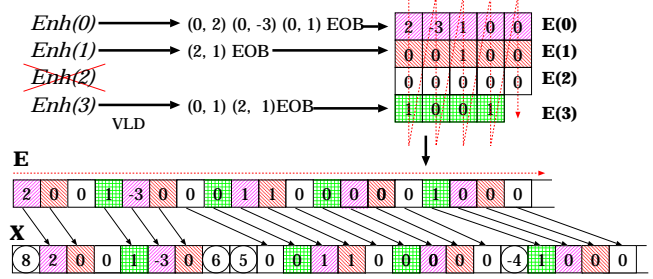


図 8 部分拡張信号損失時の合成処理

生成する [1]。さらに、 B と F を合成し X を復元する。 B を式 (9) と定義する。

$$B = \{b_0, b_1, \dots, b_i, \dots\} \quad (0 \leq i < 64) \quad (9)$$

ここで、 i はブロック内の参照順序をジグザグスキャンの順序とした際の係数位置インデックス番号を表す。また、 F を式 (10) と定義する。

$$F = \{f_0, f_1, f_2, \dots, f_j, \dots\} \quad (10)$$

B と F を合成して得られる X を式 (11) と定義する。

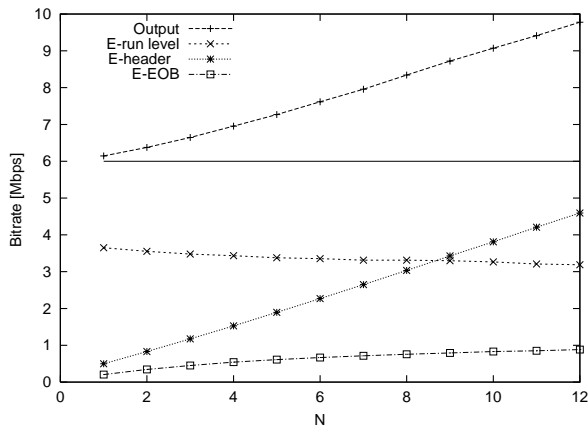
$$X = \{x_0, x_1, x_2, \dots, x_{63}\} \quad (11)$$

まず、式 (12) のように X に B を挿入する。

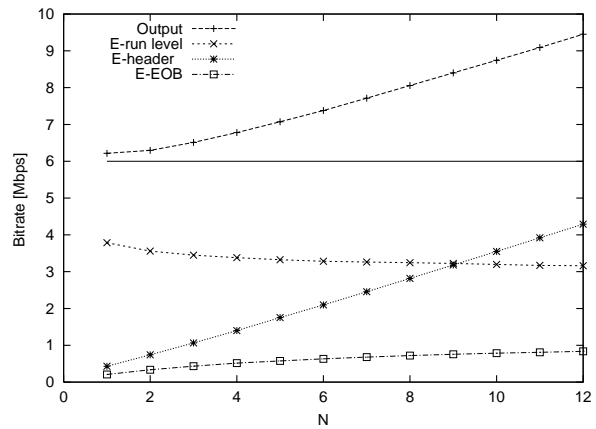
$$x_i = b_i \quad \text{for all } i \quad (12)$$

次に、 X において零係数となる x_i に順に f_j を挿入する。最後に X をスキャンし可変長符号化 (VLC) する。

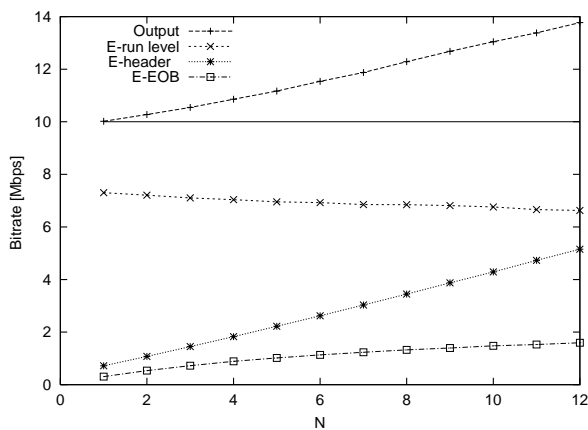
ここで、式 (7) より、 $E(n)$ を構成する各量子化係数値の F におけるインデックス番号は階層ストリーム数 N と階層ストリーム番号 n によって一意に決定される。そこで、 N と n を $Enh(n)$ のヘッダ情報に符号化することで、 $Enh(n)$ 間で互いの量子化係数値を参照する必要がなくなり、部分拡張階層ストリーム $Enh(n)$ 間の独立性が実現される。そのため、図 8 に示すように、任意の $Enh(n)$ を損失した場合においても、 $E(n) = 0$ とすることで、ストリーム構造は維持され、残りの部分拡張ストリームから映像が取得可能となる。



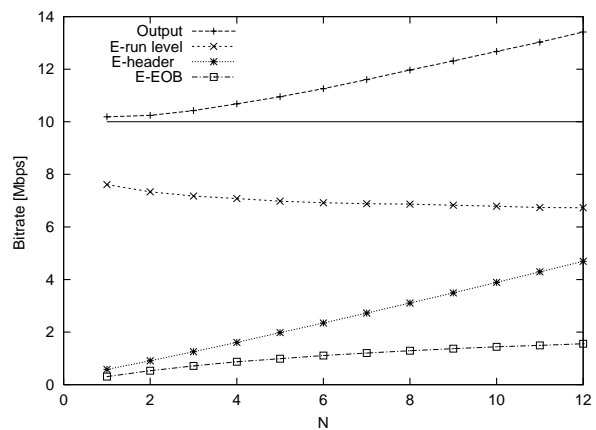
(a) Bus 6[Mbps]



(b) Tabl Tennis 6[Mbps]



(c) Bus 10[Mbps]



(d) Tabl Tennis 10[Mbps]

図 9 階層ストリーム数に対する総符号量特性

4. 評価実験

提案方式について評価を行なう．実験用ストリームの符号化条件を表 1 に示し，基本ストリームの符号化レートは 2[Mbps] とする．ただし，符号量評価においては，図 2 に示す分離器 (Separator) から出力直後の基本ストリーム *Base* と部分拡張ストリーム *Enh(n)* の総符号量を評価する．

表 1 実験環境

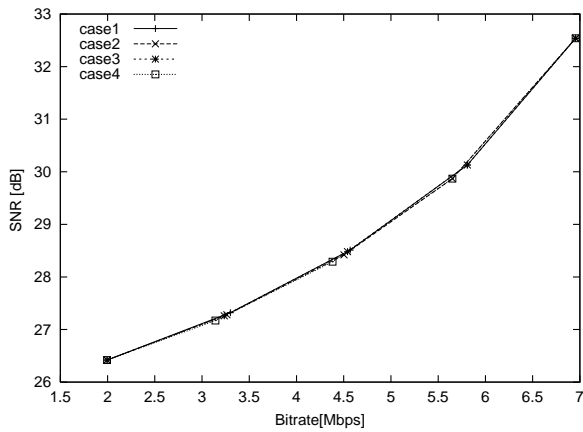
符号化方式	MPEG-2 Video MP@ML
符号化レート	6, 10 [Mbps]
符号量制御方式	TM 方式
GOP 構造	N=15, M=3 (固定)
テストシーケンス	Bus, Table Tennis
画像フォーマット	ITU-R BT.601 4:2:0 Format 輝度信号 704 pel × 480 line 色差信号 352 pel × 240 line
符号化フレーム	150 [Frame]

4.1 階層ストリーム数に対する総符号量特性

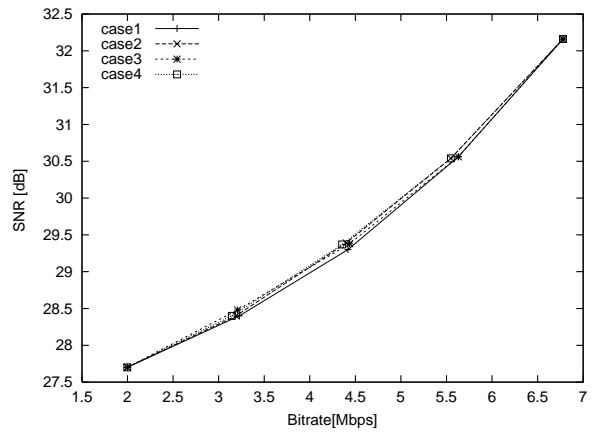
非階層ストリームを階層構造へ変換する際の総符号量特性を評価する．実験では，部分拡張ストリームを構成するヘッダ情報 (header)，係数 (run level)，係数の終了を表す EOB (End Of Block) 符号についても符号量を測定した．ただし，ヘッダー

情報はシーケンスヘッダ，ピクチャヘッダ，スライスヘッダ，MB 層データ，さらに階層ストリーム数，階層ストリーム番号によって構成される．また，階層ストリーム数 N は $1 \leq N \leq 12$ とした．実験結果を図 9 に示す．

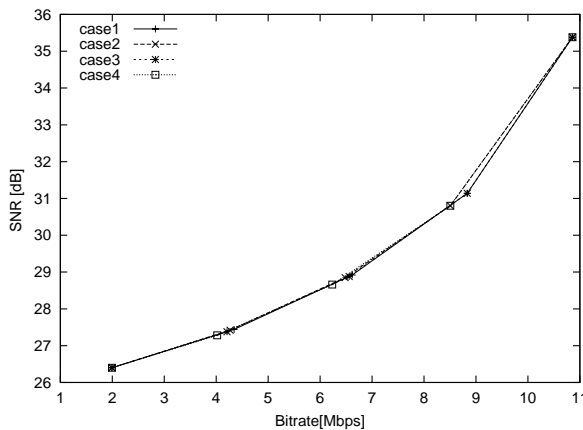
提案方式適用時において， N が小さい場合には総符号量の増加率は 1.2 倍程度に抑えられ，大幅な符号量の増大は回避されている．その理由として，提案方式ではラン長の値域が大幅に減少することが考えられる．非階層ストリームにおける値域が 0 から 63 の 64 通りであるのに対し，提案方式ではブロック内の信号を係数位置に基づき N 個に分離することから 0 から $[64/N]$ に減少する．そのため N が大きいほど 1 有意係数あたりの平均符号長が減少し，総符号量を抑制する効果が働く．長いラン長を多く含む高レートで符号化された非階層ストリームにおいて特に効果が現れる．しかし， N に比例して総符号量が増加するため N が大きい場合には課題となる．その原因として，部分拡張ストリームに含まれるヘッダ情報量が増加することがあげられる． N の増加に伴い，符号化対象となる係数ブロックの総数が増加し，MB の符号化モードに必要な符号量が加算されるためである．ただし，本実験では標準方式で規定される符号化テーブルを用いた評価に留まっており，符号化効率の最適化は施されていない．今後は，ランレングスの確率分布の定量的評価とそれに基づく符号化テーブルの最適設計により，階層ストリーム数に対する総符号量の増加率を抑制して符号化



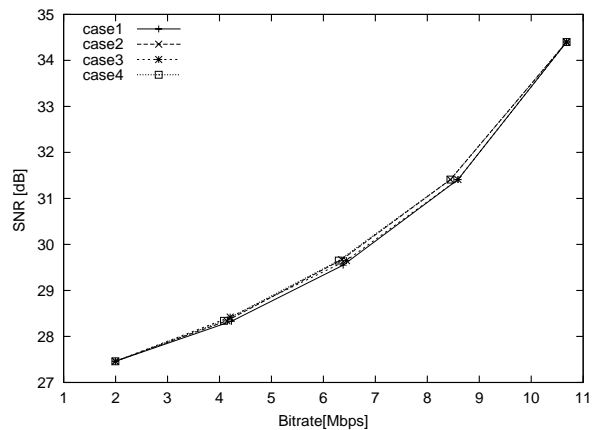
(a) Bus 6[Mbps]



(b) Table Tennis 6[Mbps]



(c) Bus 10[Mbps]



(d) Table Tennis 10[Mbps]

図 10 合成前総符号量と画質の関係

表 2 部分拡張ストリームの合成順序

	部分拡張ストリームの合成順序			
	1 番目	2 番目	3 番目	4 番目
case 1	Enh(0)	Enh(1)	Enh(2)	Enh(3)
case 2	Enh(1)	Enh(2)	Enh(3)	Enh(0)
case 3	Enh(2)	Enh(0)	Enh(1)	Enh(3)
case 4	Enh(3)	Enh(2)	Enh(1)	Enh(0)

効率を改善することが課題となる。

4.2 合成階層数と画質の関係

階層ストリーム数を $N = 4$ とし、基本ストリーム *Base* に対し、部分拡張ストリーム $Enh(n)$ を順に合成した際の合成前総符号量と画質の関係について考察する。提案方式は、任意の順序での合成を可能とするものであり、実験では、表 2 に示す 4 通りについて評価を行なう。図 10 に実験結果を示す。

実験結果より、基本ストリームと任意の順序で部分拡張ストリームを合成可能であり、合成する部分拡張ストリームの種類によらず符号量の増加、品質の向上はほぼ等しいことがわかる。しかし、本実験では、階層ストリーム数に応じた特性の比較評価は行っていない。今後は、符号化効率の関点から、アプリケーションに応じた最適な階層ストリーム数についての検討も課題である。

5. むすび

本稿では、階層型映像伝送における課題を解決するための手段として、階層間独立スケーラブル映像符号化方式を提案した。評価実験により、階層間の依存関係を排除した階層構造においても、任意の階層の組合せによる映像再生を可能とし、符号量肥大化の回避も実現可能であることを示した。

文 献

- [1] 永吉, 花村, 笠井, 富永, "MPEG-2 ビットストリーム分離・合成機能によるスケーラブル映像符号化方式," 信学論 (D)-II, vol. J84-D-II, no. 12, pp. 2525-2540, Dec. 2001.
- [2] 角田, 永吉, 花村, 富永, "ビットストリーム変換処理による多階層映像符号化方式に関する検討," 信学春季総大, 2003 年 3 月.
- [3] 涌井, 永吉, 花村, 富永, "ビットストリーム変換によるスケーラブル映像符号化方式に関する検討," PCSJ2003, P-2.01, pp. 13-14, Nov. 2003.
- [4] ISO/IEC 13182-2:2000 "Generic Coding of Moving Picture and Associated Audio, Part-2 Video," 2000.
- [5] ISO/IEC 14496-2:2001 "Generic Coding of Audio-Visual objects, Part-2 Visual," 2001.
- [6] 西村, 笠井, 高屋, 亀山, 榎, 花村, 富永, "リアルタイム MPEG-2 ビデオトランスコーダソフトウェアの開発", 情処研報, AVM25-5, pp. 25-30 (1999).
- [7] ISO/IEC 13182-1, IS "Information Technology - Generic Coding of Moving Picture and Associated Audio Information: Systems," 1995.