

## インターネット映像伝送を考慮した映像符号化方式

清水 淳 嵯峨田 淳 渡辺 裕 小寺 博

NTT ヒューマンインタフェース研究所

〒 239 横須賀市光の丘 1-1

TEL: 0468-59-2830, FAX: 0468-59-2829

Email: shimizu@nttvdt.hil.ntt.co.jp

現在、インターネット上では、その普及とともに種々の映像符号化方式を用いた映像伝送サービスが行われている、その多くは、既存の標準化された映像符号化方式を採用している。しかし、これらの方式は、低レートにおける画品質が不十分であるなど、必ずしもインターネット上での映像伝送サービスの要求事項を満たしているとは言えない。本稿では、標準符号化方式の問題点を明確にし、インターネット映像伝送サービスに適した映像符号化方式を提案する。

## Video Coding for the Internet Video Transmission

Atsushi Shimizu, Atsushi Sagata, Hiroshi Watanabe, and Hiroshi Kotera

NTT Human Interface Laboratories

1-1 Hikarinooka, Yokosuka, Kanagawa 239, JAPAN

TEL: +81-468-59-2830, FAX: +81-468-59-2829

Email: shimizu@nttvdt.hil.ntt.co.jp

Video transmission services on the Internet have attracted the attention of many people. Most of the video coding method adapt the conventional video coding standard. However, those standards are not necessarily fit to the Internet environment. In this paper, a new video coding method using global motion compensation(GMC) for the Internet video transmission is proposed.

## 1 はじめに

近年、インターネットの利用が急速に拡大している。このインターネットを通じて、WWWに代表される様々なサービスが提供されている。その中で、インターネットを使った映像伝送サービスが注目を集めている。現在、インターネットを使った映像配信などが盛んに行なわれている。映像伝送サービスが拡大した背景には、パソコンの高性能/高機能化により、従来、専用ハードウェアが必要であったリアルタイム復号処理がソフトウェアで可能となった点が挙げられる。

現在、インターネット上で行なわれている映像伝送サービスには、様々な映像符号化方式が用いられている。その多くは、既存の標準化された映像符号化方式(H.261[1], MPEG1[2], H.263[3])を採用している。これらの標準方式を採用するメリットは、既にリアルタイムエンコーダが市販されており、デコーダソフトがソース形式で公開されているため、専門家でなくとも容易に符号化技術を利用できることにある。

しかし、これら標準方式が、必ずしもインターネット映像サービスに求められる性能や機能を満たしているとは言い難い。それは、インターネット映像サービスの要求事項が、各標準方式の制定時に想定していたアプリケーションと異なるためである。

また、インターネットには、ユーザにより伝送レートが異なるという問題があり、異なる符号化レートのビットストリームを複数用意する必要がある。この要求事項を満足させる方法として、複数のエンコーダによるビットストリーム生成やスケラビリティ機能[4]を持つ符号化方式の採用、再符号化による符号化レートの変換[5]などが提案されている。先に、筆者らは再符号化による符号化レートの変換方法[6]を提案し、その有効性を示した。

本稿では、インターネット映像伝送における標準方式の問題点を明確にし、特にフレーム間予測効率の改善を中心に、インターネット映像伝送サービスに適した映像符号化方式を提案する。

## 2 映像符号化標準方式の問題点

ユーザがインターネットに接続する場合、現在はまだ電話回線を利用するケースが多い。その時の伝送レートは、モデムの使用により最大で約30kbps程度である。また、最近では、ISDN(64kbps以上)への移行も進んでいるが、まだ一

般的であるとは言えない。従って、インターネット上の映像伝送サービスは、低レートでの符号化効率が問題となってくる。ここでは、主に、低レートでの符号化効率の観点からMPEG1とH.263について問題点を挙げる。

### 2.1 MPEG1

MPEG1は、主に蓄積メディア用として制定された映像符号化方式である。そのビットレートは主に1Mbps以上を想定している。以下に、MPEG1を低レートで使用する場合の問題点を示す。

- ヘッダなどのオーバーヘッドのサイズ  
高機能化の為にヘッダ部に様々な情報が付加されている。符号化レートが1Mbps以上と高い場合、その影響は小さいが、64kbps以下の低レートでは、その影響を無視できない。
- 固定化されたフレームレート  
29.97Hzや25Hzなど、いくつかの決まったフレームレートにのみ対応している。このため、フレームスキップは、ピクチャヘッダやスライスヘッダなどヘッダだけの情報を送ることで擬似的に再現している。64kbps以下の低レートでは、入力映像の全フレームを符号化することは難しく、大量にフレームスキップが発生する。従って、フレームスキップのためのオーバーヘッドが大きくなる。

### 2.2 H.263

64kbps以下の低レートを考慮した映像符号化方式である。MPEG1と異なり、フレームスキップも可能である。しかし、そのアプリケーションは、テレビ電話などの映像通信分野を想定している。以下に、H.263の問題点を示す。

- 狭い動きベクトルの表現範囲  
テレビ電話やテレビ会議の場合、映像の動きが少ないため、短い動きベクトルで十分である。しかし、動きが大きい映像では、動き補償効率が低下する。特に、符号化フレームレートが低下する低レートでは、フレーム間距離が長くなることにより、動き補償効率の低下は顕著に現れる。
- 低いイントラフレームの符号化効率  
MPEG1やMPEG2では、イントラマクロブロックのDCT係数のうち、DC成分を予

測符号化することで符号化効率を向上させている。H.263では、DC成分の予測符号化はサポートされていない。

- 限定された画像サイズ  
 予め使用できる画像サイズが決まっているため、規定されている画像サイズ以外は使用できない。パソコンのビデオキャプチャ機能を利用する場合、640×480(NTSCの場合)が基本サイズとなっているため、CIFやQCIFの画像生成が複雑になる。

### 3 インターネット映像伝送に適した映像符号化方式

提案方式では、先に述べた問題を解決するために種々の改良を行っている [7][8]。符号化アルゴリズムは、ブロックベース(16×16)の動き補償とDCTを基本としている、ここでは、動き補償予測効率の改善方法を示す。

#### 3.1 グローバル動き補償

標準方式に用いられている動き補償方式(Local Motion Compensation: LMC)は、平行移動モデルに基づいてブロック単位に動きベクトル(Local Motion Vector: LMV)を伝送する。パンやチルトは、従来の動き補償方式で補償可能であるが、動き量が大きい場合は、動きベクトルが大きくなり、符号化効率が低下する問題がある。また、平行移動以外のズームなどは、補償することが不可能である。提案方式では、動き補償の予測効率を高めるため、パン、チルト、ズームなどの画面全体の動きをフレーム毎に大域動きベクトル(Global Motion Vector: GMV)として与え、動き補償を行なうグローバル動き補償(Global Motion Compensation: GMC)[9]を採用している。GMCモードのブロックは、大域動きベクトルを補正するために、LMC同様にブロック単位に動きベクトルを伝送している。GMCにより、予測効率を落とすことなく、小さい動きベクトルが使用できる。

#### 3.2 動き補償モード選択

複数の動き補償モードを切り替える符号化方式において、動き補償モードの選択には注意を要する。予測効率と動きベクトルやモード選択のためのオーバーヘッドは、トレードオフの関係にあるため、動き補償モードの選択方法によっては、予測

表 1: 各動き補償モードの比較

動き補償モード	予測効率	LMVの数
16×16 動き補償 OFF	×	0
16×16 LMC	△	1
8×8 LMC	○	4
16×16 GMCのみ	×	0
16×16 GMC+LMV	△	1
8×8 GMC+LMV	○	4

効率の低下やオーバーヘッドの増大を招く。提案方式の動き補償は、表 1に示す6種類のモードからなる。動きベクトル数は、オーバーヘッドの大きさを示している。オーバーヘッドには、動きベクトルの他にブロックの符号化モードを示す情報があるが動きベクトルの情報量に比べて小さい。表 1の中で、“16×16 動き補償 OFF”モードは、動きベクトルが(0,0)での動き補償と等価である。また、“16×16GMCのみ”モードは、LMVによる動きベクトルの補正を行なわない。

各動き補償モードは、予測効率と動きベクトルなどのオーバーヘッドに違いがある。従来方法では、各動き補償モードで算出される予測誤差を、オーバーヘッドの大きさに応じて定めた定数を使って補正していた。しかし、常に最適な動き補償モードが選択されるとは限らない問題がある。そこで、提案方式では、符号化対象ブロックの量子化ステップを用いて適応的に動き補償モードを選択している。予測誤差を差分絶対値和とする場合の動き補償モード選択の手順は以下の通りである。

1. 各動き補償モードの予測誤差  $Err_i (i=1,2,\dots,6)$  を算出する。
2. 予測誤差  $Err_i$  から最小予測誤差  $Err_{min}$  を算出する。
3. 符号化対象小ブロックの量子化ステップ  $Q$  を算出する。
4. 量子化ステップ  $Q$  から許容値  $T$  を次式より算出する。  

$$T = C \cdot Q \quad (1)$$
 ここで、 $C$  は定数である。
5. 各動き補償モードの予測誤差  $Err_i$  で、次式を満たす動き補償モードを候補とする (図 1(a))。

$$Err_i \leq Err_{min} + T \quad (2)$$

6. 動き補償モードの候補の中で、最もオーバーヘッドが小さい動き補償モードを、符号化対象ブロックの動き補償モードとする (図 1(b))。

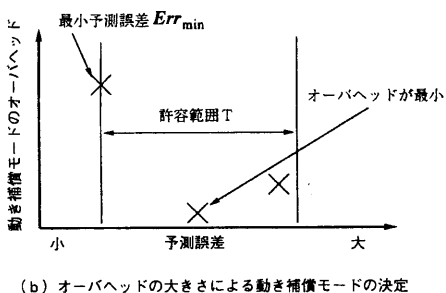
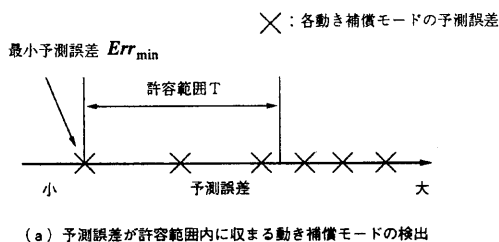


図 1: 動き補償モードの選択方法

## 4 シミュレーション

### 4.1 シミュレーション条件

GMC と動き補償適応切り替えの効果を見るため、以下の三方式の性能を比較した。

1. H.263 TMN5[10]
2. 提案方式 (a) (従来の MC モード選択)
3. 提案方式 (b) (適応 MC モード選択)

比較対象として H.263 を用いた理由は、フレームスキップに対応でき、MPEG1 より低レートでの映像伝送に適しているためである。

表 2 にシミュレーション条件を示す。H.263 は、算術符号を用い符号化効率を高めている。提案方式 (a),(b) の違いは、動き補償モードの選択方法のみである。また、今回は、符号化効率の比較のため、符号化フレームレートを 15fps とし、量子化パラメータは固定とし、1 から 31 まで変化させた。LMV の探索範囲は  $\pm 15 \times \pm 15$  とした。動き補償モード選択に用いた予測誤差の指標は、次式に示す差分絶対値和を用いた。

$$SAD = \sum_{j=0}^{15} \sum_{k=0}^{15} |P_c(j, k) - P_r(j, k)| \quad (3)$$

ここで、 $P_c(j, k)$  は符号化対象ブロックの画素、 $P_r(j, k)$  は参照画像のブロックの画素である。

テストシーケンスには、MPEG4 で用いられている “Foreman”, “Coast guard”, “News” の 3 種類を用いた。 “Foreman” と “Coast guard” は、パンやチルトなどの画面全体の動きがあり、 “News” は、画面全体の動きが小さい画像である。画像サイズは QCIF (176x144)、フレームレートは 30fps、入力フレームの総数は、300 フレームである。符号化フレームレートを 15fps としているので、符号化フレームの総数は、150 フレームとなる。

### 4.2 実験結果

図 2 に、各方式の符号化性能を示す。横軸は 1 フレーム当たりの平均発生情報量、縦軸は符号化フレームの平均 PSNR である。図 2 に示すように、H.263 に比べ、各画像とも提案方式の性能が高いことがわかる。特に、H.263 と提案方式で、平均発生情報量が大きくなるに従い差が拡大している。これは、提案方式が GMC より動き補償効率が向上し、予測誤差が減少したためと考えられる。特に、 “Foreman” と “Coast guard” は、画面全体の動きがあり、GMC によるゲインが大きいのと思われる。これに対し、 “News” は、画面全体の動きは小さく、背景も静止しているため、GMC の効果は期待できない。さらに、GMV や GMC モード選択のための情報がオーバーヘッドになる可能性がある。そのため、H.263 との差は、 “Foreman” と “Coast guard” に比べ小さくなっている。

図 3 に、動き補償モードの選択方法による符号化効率の比較を示す。横軸は量子化パラメータ、縦軸は提案方式 (b) (適応動き補償モード選択方法) による提案方式 (a) (従来の動き補償モード選択方法) に対するゲインを示している。ゲインは次式により算出した。

$$Gain = \frac{G_a - G_b}{G_a} \quad (4)$$

ここで、 $G_a$  は提案方式 (a) の発生情報量、 $G_b$  は提案方式 (b) の発生情報量である。

図 3 から分かるように、量子化パラメータにより、動き補償モード選択によるゲインは異なっている。このことは、量子化パラメータ毎に、最適な予測誤差の補正値が存在することを示している。傾向として、量子化パラメータが大きくなるに従い、量子化パラメータに基づく動き補償モード選択の効果が大きくなっている。量子化パラメータが大きい場合、動き補償によるゲインに比

表 2: シミュレーション条件

	H.263	Method (a)	Method (b)
Global MC	OFF	ON	ON
Adaptive MC Mode Selection	OFF	OFF	ON
Intra DC Prediction	OFF	ON	ON
Syntax-based Arithmetic Coding	ON	OFF	OFF

べ、動きベクトルなどのオーバーヘッドが問題になるためである。

#### 4.3 考察

以上に示したように、GMCはカメラ操作により画面全体の動きがある場合に有効である。H.263で表現できる動きベクトルを拡張した場合、平行移動に関して動き補償の予測効率はGMCと同等になると考えられる。しかし、動きベクトルの長さが長くなることで、オーバーヘッドが大きくなる。特に、低レートでは、動きベクトルなどのオーバーヘッドの影響が無視できないため、インターネット映像伝送には適さないと考えられる。

さらに、動き補償モードの選択に量子化パラメータを用いることで、動き補償予測効率を落とさずに、動き補償のオーバーヘッドを削減できることが分かった。一定の量子化パラメータで符号化を行う場合、予測誤差の補正値の最適値求めて固定にすることも可能であるが、一定の符号化レートで符号化する場合、量子化パラメータは変動するため、適応的な動き補償モードは有効である。

また、イントラブロックのDC成分予測符号化と、機能限定によるヘッダ部のオーバーヘッド削減も符号化効率改善に寄与している。

#### 5 むすび

インターネットでの映像伝送におけるMPEG1やH.263の問題点を明らかにした。これらの問題のうち、動き補償予測効率の改善を試みた。この手法を用いて、インターネット映像伝送に適した映像符号化方式を提案し、H.263に対する優位性を検証した。シミュレーションにより、提案方式に採用したグローバル動き補償と動き補償モードの適応選択手法の有効性を確認した。

#### 参考文献

[1] ITU-T: "Video codec for audiovisual services at p x 64kbit/s", Recommendation H.261, 1990.

[2] ISO/IEC: "Coding of moving picture and associated audio for digital storage media up to about 1.5Mbit/s", IS-11172, 1993.

[3] ITU-T: "Video Coding for Low Bitrate Communication", Recommendation H.263, 1995.

[4] ISO/IEC: "Information Technology- Generic Coding of Moving Picture and Associated Audio", IS-13818, 1994.

[5] H. Sun, W. Kwok, J.W. Zdepski, "Architectures for MPEG Compressed Bitstream Scaling", IEEE Trans. Circuits Syst. Video Technol., vol. 6, no. 2, pp.191-199, Apr. 1996.

[6] 清水, 嵯峨田, 渡辺, 小寺: "フレームスキップによるビットストリームスケールリング", PCSJ'96, pp.25-26, 1996.

[7] 嵯峨田, 清水, 渡辺, 小寺: "インターネット映像伝送を考慮した映像符号化の効率改善", 信学総大'97, 1997.

[8] 清水, 嵯峨田, 渡辺, 小寺: "インターネット映像伝送を考慮した映像符号化方式", 信学総大'97, 1997.

[9] 上倉, 渡辺, "動画像符号化におけるグローバル動き補償法", 信学論 B-I, Vol. J76-B-1, No. 12, pp. 944-952, 1993-10.

[10] Karl O. Lillevold, "Telenor H.263 encoder: tmn Version 2.0," <ftp://bonde.nta.no/pub/tmn/software/>, 1996.

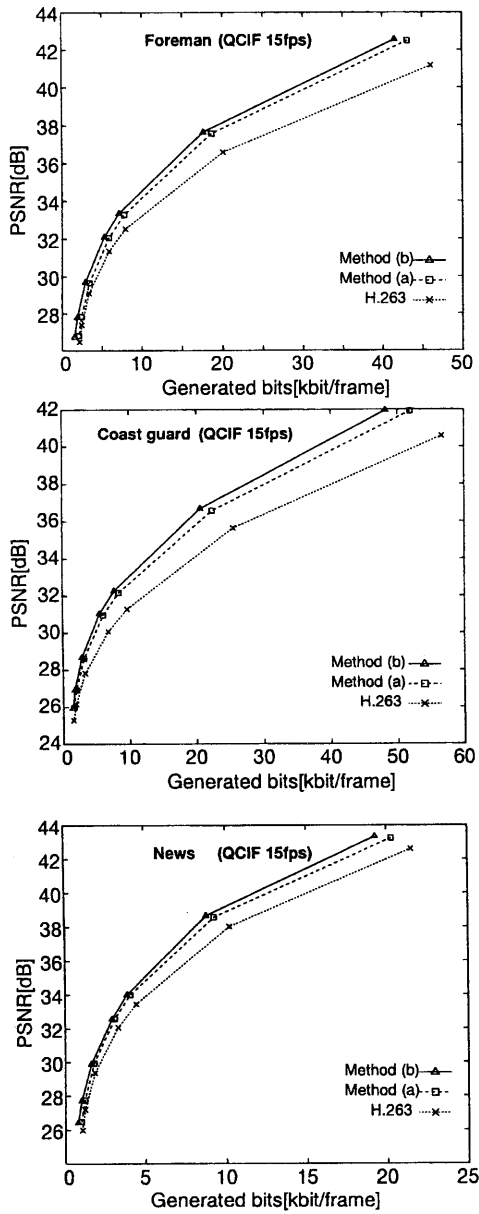


図 2: 符号化性能

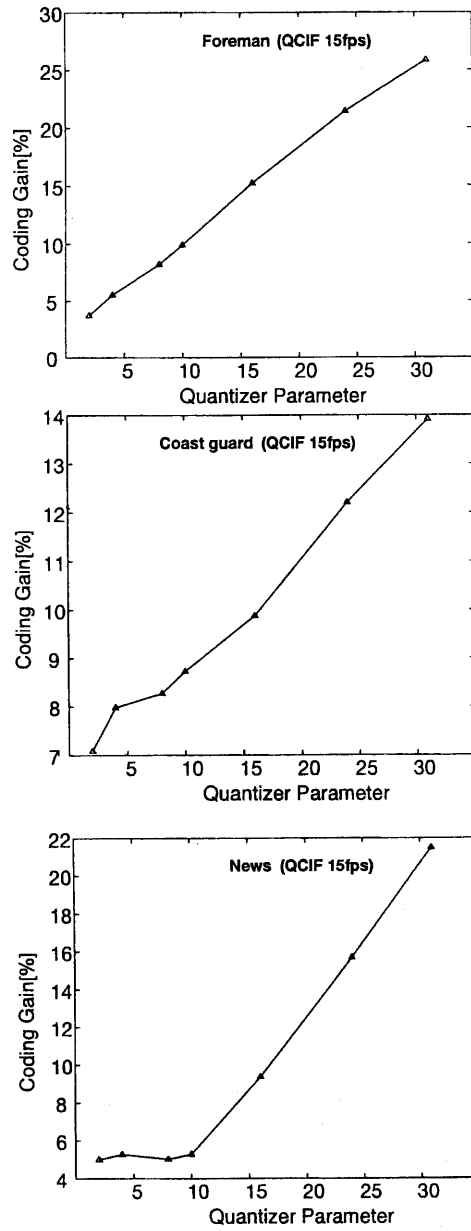


図 3: 適応 MC モード選択による利得