

高度前処理技術の適用による MPEG-4 の画質改善に関する一検討*

特別-5

宮地 悟史†

松本 修一

株式会社 KDD 研究所 画像通信グループ‡

1 はじめに

最新の画像符号化国際標準方式 MPEG-4 が勧告化され、インターネットビデオ通信分野における標準的な符号化方式として期待されている。インターネットビデオでは、ユーザがインターネットへアクセスするための回線容量が、モデムや PHS といった数十 kbit/s が一般的であるため、これらに合わせた超低ビットレートでの符号化が行われる。

MPEG-4 は自然画像、人工画像、静止テクスチャといった様々なプロファイルを有しており、対応ビットレートも幅広い。その中で今後も超低ビットレートで用いられると予想される MPEG-4[1] 自然画像の符号化アルゴリズムに関しては、従来最も一般的であった ITU-T H.263[3] と比較して、演算面での細かな工夫に加えてシンタックスの見直しや VLC の改良等により符号化効率の面でいくつかの向上が図られている。

筆者らは文献 [4] において、高度符号化前処理を利用した統合的符号化制御を提案し、低ビットレート符号化に有効であることを示した。本稿では、MPEG-4 のさらなる画質向上を目的として、統合的符号化制御を MPEG-4 に適用するための検討を行う。

2 高度前処理技術の MPEG-4 への適用

本方式では、図 1 に示す前処理部を、MPEG-4 VM[2] に適用する。

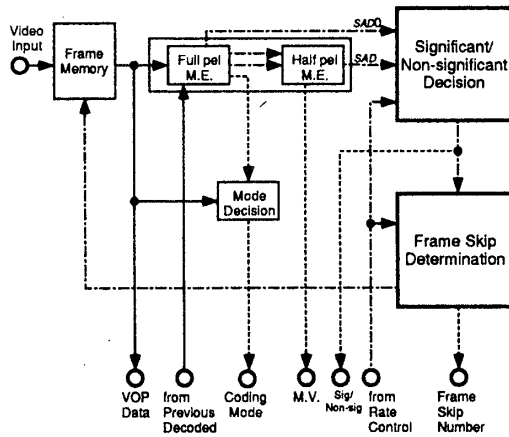


図 1：前処理部のブロック構成図

前処理部では、

- 入力フレームのスキップ
- 動き検出
- 符号化モード (Intra/Inter) 判定
- マクロブロック有意判定 (空間領域情報量制御)
- フレームスキップ数決定 (時間領域情報量制御)

の各処理を行う。通常の MPEG-4 VM で行われている動き検出に加えて、空間領域情報量制御としてマクロブロック単位の有意/非有意判定、および時間領域情報量制御としてフレームスキップ数の決定を行う。

前処理の結果、時空間制御情報および動きベクトル、符号化モードの各情報が符号化部へ入力される。符号化部では、前処理部で得られた情報を量子化パラメータ制御に反映させ、振幅領域情報量制御を行う。これら時空間振幅領域の統合制御により、MPEG-4 シンタックスに従い符号化を行う。

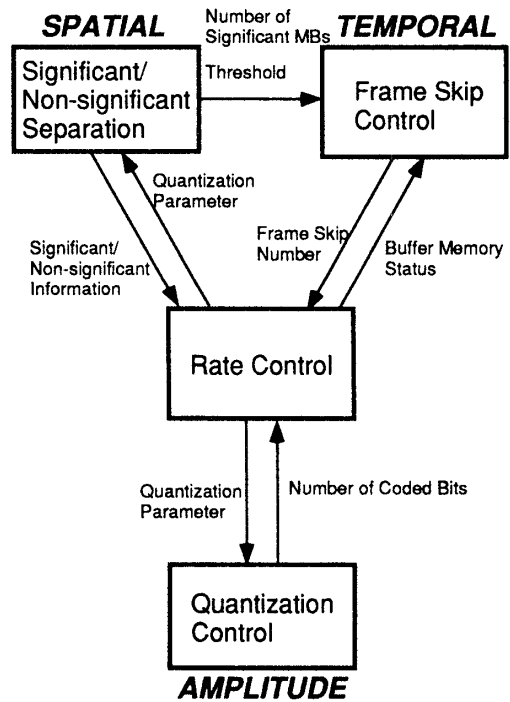


図 2：統合制御のメカニズム

以下に前処理部の各制御の詳細について述べる。

2.1 空間領域情報量制御

空間領域情報量制御は、符号化前処理における有意/非有意マクロブロック分離により行われる。マクロブロック毎に動き検出を行い、動き予測誤差を算出する。動きベクトル (0, 0) における予測誤差絶対値と SAD_0 がしきい値 T_H を下回る場合は、静止ブロックとして取り扱い、そのマクロブロックは not coded ブロックとして符号化をスキップされる。また、通常の動き予測誤差絶対値と SAD が閾値 T_H を下回った場合、非有意ブロックと判定して予測誤差の量子化を一切行わず、動きベクトルのみを伝送することとする。

このように分類されたマクロブロックは、表 1 の通り伝送が行われる。空間領域情報量制御は、

1. 量子化により切り捨てられる係数の除去による有意領域の特定
2. 有意ブロック数からの発生情報量の予測

を目的として行われ、これらの目的が満たされるよう判定しきい値 T_H が決定される。したがって、しきい値 T_H は量

*Quality improvement of MPEG-4 video using with advanced pre-coding technique

†Satoshi MIYAJI and Shuichi MATSUMOTO

‡KDD R&D Laboratories, Inc. Visual Communications Group

子化パラメータに応じて、量子化により切り捨てられるべき係数が非有意となるよう設定される。このとき、有意ブロックには情報の発生する係数のみが残されることとなり、有意ブロックの個数からフレームの情報発生量を予測することが可能となる。

表1: マクロブロックの分類

判定	判定条件	符号化伝送情報
静止	$SAD_0 < T_H$	MB スキップ (<i>not_coded=1</i>)
非有意	$SAD < T_H$	動きベクトル, <i>cbp=0</i>
有意	$SAD \geq T_H$	通常符号化

2.2 時間領域情報量制御

時間領域情報量制御では、フレームレートの決定を行う。MPEG-4のシンタックスでは、前フレームとの時間間隔を示す *VOP_time_increment* により、可変フレームレート符号化が可能となっている。しかしながら、MPEG-4 VM では固定のフレームレートを基本としており、フレームの符号化後に発生した情報量がバッファ残量を圧迫する場合にフレームスキップを行うといった手法がとられているため、

- 画像性質を反映したフレームレート制御が行えない。
- リアルタイム伝送においては、フレームスキップが伝送遅延量となるため、符号化後の事後決定では最大遅延量の規定ができない。

といった問題がある。

本統合的符号化制御では、空間領域情報量制御であるフレーム内の有意ブロック数からフレームの持つ情報を予測し、かつバッファメモリ占有量を考慮して、フレーム間隔の決定を行う。フレーム間隔を符号化前に決定することで、要求される画像品質やフレームレートに応じた情報量の割り当てや、最大フレーム遅延量の規定が可能となる。バッファメモリに余裕のあるときには、画像品質を保ちつつ低遅延伝送となるよう制御を行い、バッファが上限に近いときには画像破綻の防止と伝送遅延量の極端な増大が生じないようにフレームレート制御を行う。

2.3 振幅領域情報量制御

本方式3つ目の制御である振幅領域情報量制御では、与えられた量子化パラメータにより MPEG-4 VM の量子化アルゴリズムにしたがって量子化を行い、その結果発生する情報量をレート制御に反映させる。

実際の量子化パラメータの決定は後述のレート制御で行われ、前述の時間/空間情報量制御によるビット割り当てと、実際の発生情報量との相違を吸収するよう行われる。

2.4 レート制御

レート制御は、MPEG-4 VM の方式を基本とし、VM では平均値をベースに算出される目標符号化ビット数を、本方式では前処理部により得られる予測値を利用して、より精度の高いレート制御を行う。

MPEG-4 VM では、以下のモデル関数によるレート制御が用いられており、

$$R = X1 \frac{S}{Q} + X2 \frac{S}{Q^2} \quad (1)$$

ここで、 R : 符号化ビット数、 S : 画像の複雑さ (フレーム間差分絶対値和 SAD)、 $X1, X2$: モデリング係数である。上式のモデルを適用し、二次方程式の解として量子化パラメータを算出する。

本方式では、前処理による時空間情報量制御により符号化前にフレームレートが決定され、このフレーム割り当て情報量を、式(1)の目標符号化ビット数 R として用いる。また、符号化の難易度を推定するためのフレーム間差分絶対値の算

出も、有意領域に対してのみ行い、実際の発生量との相関を高めている。

3 性能検証実験

これまでで述べた符号化前処理を実際の MPEG-4 VM に適用し、計算機シミュレーションにより性能検証を行った。表2に本方式と MPEG-4 VM のシーケンス全体の SNR、平均フレームレート FR 、および SNR の変化量 ΔSNR をそれぞれ示す。符号化ビットレートは、最も普及していると考えられるインターネットアクセス回線の V.34 モデム (33.6kbit/s、実効 25kbit/s 程度) および PIAFS (29.2kbit/s) において、音声に相当する 5~6kbit/s およびその他のオーバーヘッド情報を除いた 18.4kbit/s に設定した。

表2: 実験結果 (at 18.4kbit/s)

	MPEG-4 VM		本方式		ΔSNR (dB)
	SNR (dB)	FR (fps)	SNR (dB)	FR (fps)	
画像					
carphone	28.8	5.8	29.3	9.0	+0.5
foreman	27.5	6.9	28.0	6.9	+0.5
news	26.5	5.1	27.4	6.5	+0.9

表2の結果から、SNRの点で本方式が MPEG-4 VM の特性を上回っていることがわかる。また、carphone、news についてはフレームレートの点においても本方式の方が高く、このことを考慮すると実質の SNR 改善量は表2の ΔSNR の値以上のものとなっている。これは、本方式が有意な領域に対して最適なビット配分を行い、それに基づき最適なフレームスキップを符号化前に与えているのに対し、MPEG-4 VM が、全ての領域に対して符号化を行い、符号化後のバッファ残量に依存してフレームスキップ制御を行っており、非有意領域に対する不必要な情報発生と、この情報発生による過剰なフレームスキップが行われているためである。

本方式による特性改善は、テレビ放送画像である news が最も大きくなっているが、このニュース報道番組は固定されていないカメラで取材対象を追いかけるといった画面全体が大きく変化する画像であり、さらにフレーム毎の画像性質の変化が大きいため、符号化前の画像分析が大きな効果をもたらす結果となった。

4 まとめ

標準的な超低ビットレート画像符号化方式 MPEG-4 VM に対し、統合的時空間符号化制御方式を適用する検討を行った。その結果、MPEG-4 のシンタックスに準拠した上で、通常の VM シミュレーションと比較して画像品質が改善され、本方式の MPEG-4 への有効性が確認された。

今後の課題として、ソフトウェアによる MPEG-4 リアルタイムエンコーディングの検討が挙げられる。

日頃御指導頂く (株) KDD 研究所の村谷代表取締役所長に感謝いたします。

参考文献

- [1] ISO/IEC FDIS 14496-2: "Information Technology - Coding of Audio Visual Objects - Part 2: Visual" (Oct. 1998)
- [2] ISO/IEC N2552: "MPEG-4 Video Verification Model Version 12.0" (Mar. 1998)
- [3] ITU-T Recommendation H.263: "Video Coding for Low Bit Rate Communication" (Mar. 1996)
- [4] 宮地, 松本: "統合的時空間制御による超低ビットレート動画像符号化方式", 映情学誌, Vol.51, No.10, pp.1706-1714 (Oct. 1997)