

低レート向け動画像符号化方式

New Video Encoding Algorithm for Low-bit-rate Applications

あらまし

PHSのデータ通信サービスや、既存のアナログ電話回線用の高速モデムを用いれば、数十kbpsのデジタルデータ通信を行うことは容易であり、これを利用すれば、動画像通信を行うこともできる。ところで、ITU-Tで標準化された動画像符号化方式であるH.263は、様々な画像を少ない情報量に圧縮できる優れた方式である。しかし、大きな動きが含まれる画像では多くの情報が発生し、伝送できる画面の数が少なくなってギクシャクした動きになったり、各画面の品質が低下したりすることは避けられない。そこで、これらの画質低下要因を取り除き、より高い品質の画像を提供するために、解像度を各画面の状況に依存して適応的に切り替える方式を考案した。本方式は、H.263の改良版であるバージョン2に採用された。

Abstract

High-speed (from 30 to 64 kbps) digital data transmission services have become available through the personal handy phone systems (PHS) and high-speed modems communicating over existing analog telephone lines. These services can transmit motion pictures by using video coding algorithms.

H.263, which was standardized by the International Telecommunication Union-Telecommunication Standardization Sector (ITU-T), is a very-powerful video coding algorithm with high compression capability for various types of video sequences. However, very dynamic scenes suffer from frame skipping or picture quality degradation because of the large amount of data they contain. To minimize these effects, we proposed an algorithm for adjusting the screen resolution to suit the picture conditions. This algorithm has been adopted in ITU-T H.263 Version 2.



伊藤 隆 (いとう たかし)

1983年東京大学大学院情報工学修士課程了。同年(株)富士通研究所入社。以来動画像符号化方式、装置、システムの開発に従事。昭和63年度電子情報通信学会藤原記念学術奨励賞受賞。
マルチメディアシステム研究所メディア処理研究部



中川 章 (なかがわ あきら)

1991年東京大学大学院電子工学科修士課程了。同年(株)富士通研究所入社。以来動画像符号化装置および方式開発に従事。1997年MoMuC 97優秀論文賞受賞。
マルチメディアシステム研究所メディア処理研究部



森松映史 (もりまつ えいし)

1987年東京工業大学大学院総合理工学研究科物理情報工学専攻修士課程了。同年(株)富士通研究所入社。以来動画像符号化方式および装置の開発に従事。
マルチメディアシステム研究所メディア処理研究部

ま え が き

ISDNというまでもなく、PHSでも32 kbps(ユーザが使えるのは29.2 kbps)のデータ通信サービスが提供され、さらなる高速化の動きもある。従来からのアナログ電話網でも、高速モデムの普及により数十kbpsのデジタルデータの送受信が容易にできる環境にある。これらの通信インフラを利用するサービスの一つとして、動画像通信が挙げられる。しかし動画像は100 Mbpsを超える多量の情報量を持つため、高性能な動画像符号化手法の導入が必須である。

ところで、ITU-Tで国際標準化されたH.263は、最新の研究成果を取り入れた優れた動画像符号化技術で、多様な画像を効率的に符号化することができる。しかし、大きな動きを含むシーンでは大量の情報の発生は避けられず、通信速度一定の条件の下では、単位時間に送信できる画面の数が減少し、再生される動きがギクシャクしたものになる、あるいは各画面の品質が下がるといった、品質の低下が生じる。これはH.263特有の現象ではなく、それより前に標準化されたH.261、あるいはISO/IECで標準化されているMPEG-1、-2、-4といった方式が、いずれもフレーム間予測を採用していることによる。動きが大きいシーンでは、画面間の相関が高いというフレーム間予測の前提が満たされない。

そこで、このような画質低下要因を取り除くために、フレーム間予測方式で伝送される予測誤差情報の解像度を、画面ごとに適応的に切り替える方式を考案した。

予測誤差解像度適応選択方式

● 原理

動画像の主要な高能率符号化方式では、画面間の相関

が強いことを利用して情報量の削減を実現している。これをフレーム間予測符号化方式と呼び、本方式もこの範ちゅうに入る。図-1を参照しながら動作原理を示す。

- ・送信側で、送ろうとする画面と参照する画面とから、変化した部分を抽出する。
 - ・この変化部分を送るのに要する情報量が多いと判断したときは、解像度を下げる(図中①)。
 - ・縮小した変化部分を伝送する(図中②)。
 - ・受信側では、受け取った縮小変化部分を拡大する(図中③)。
 - ・拡大した変化部分を、送信側と共通の情報である参照画面に加えることで、送られた画面を得る(図中④)。
- 本方式が有効に機能することは、以下の2点に負っている。
- ・解像度を下げて情報量を抑えなければいけないのは、変化部分の面積が大きいときである。
 - ・そのような動きを人の目は鮮明には捉えられない。

● 符号器構成

本方式をブロック構成図として図-2に示す。本稿では詳細な説明⁽¹⁾は割愛し、考案方式に特有な部分の概要を以下に述べる。

● 解像度判定部

予測誤差信号の解像度を決定する中心的機能を持つブロックであり、画面ごとの「アクティビティ」に従って判定を行う。すなわち、 $\overline{QP}_{i-1}$ を前の画面で用いられた量子化パラメタ QP の平均値(言い換えれば画面のきれいさ)で、値が小さいほど高品質な再生画像が得られるパラメタ、また B_{i-1} を前の画面の発生情報量として、以下のアルゴリズムを実行する。

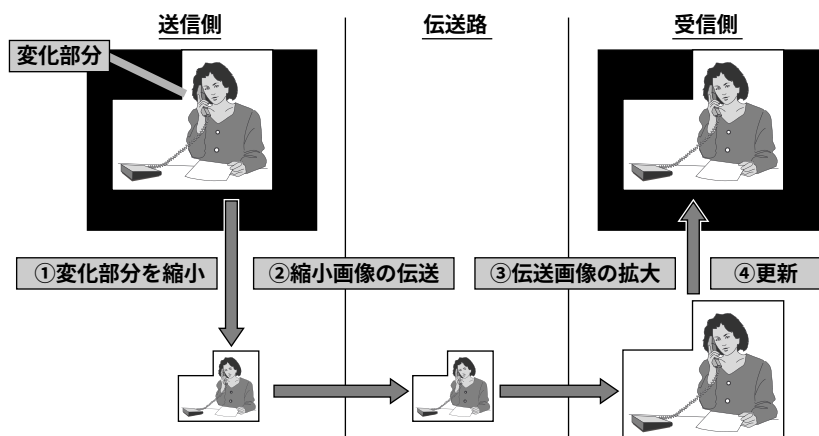


図-1 考案方式の原理
Fig.1-Basic idea of proposed algorithm.

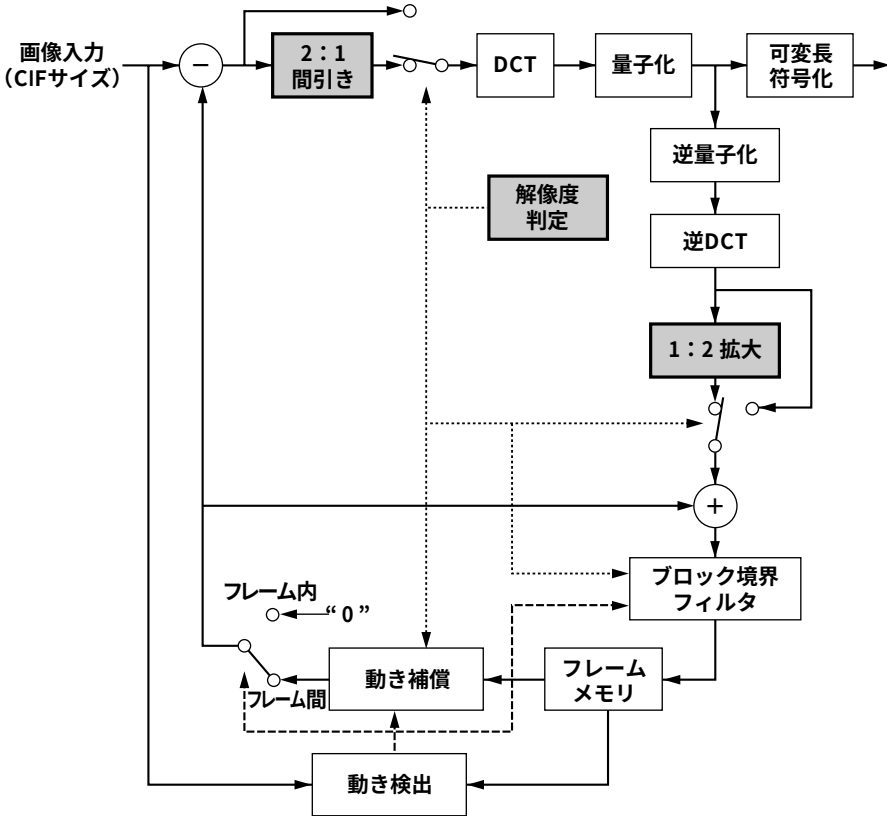


図-2 考案方式のブロック構成
Fig.2-Block diagram of proposed algorithm.

【通常解像度から縮小解像度への切り替え】

$$\text{if } (\overline{QP_{i+1}} \times B_{i+1} > TH1) \quad \dots \quad (1)$$

Switch to reduced-resolution ;

この意味は、画質が下がる方向に量子化パラメタを大きな値にしても多くの情報量が発生したときは、解像度を下げることである。

なお、閾値 THI は、通常解像度で再生品質の点から使用限界となる一番粗い(大きな値の)量子化パラメタ「 $QP1$ 」と1秒あたりの画面数(フレームレートという)「 $FR1$ 」とから、

$$TH1 = QP1 \times (\text{設定ビットレート} / FR1) \cdots (2)$$

により計算する。

【縮小解像度から通常解像度への切り替え】

$$\text{if } (\overline{QP_{i-1}} \times B_{i-1} < TH2) \quad \dots \quad (3)$$

Switch to normal resolution ;

画質が上がる方向に量子化パラメタを小さな値にしても情報が少ししか発生しないときは、解像度を上げるという判定を行う。*TH2*は、縮小解像度で再生品質の点から使用限界となる一番良い(小さな値の)量子化パラメタ「*QP2*」とフレームレート「*FR2*」とから、

$$TH2 = OP2 \times (\text{設定ビットレート} / FR2) \cdots (4)$$

で求める。

後に結果を示すシミュレーションでは、 $QP1 = 16$, $FR1 = 6$, $OP2 = 7$, $FR2 = 8$ とした。

【縮小から通常への解像度切り替え時の係数制限】

縮小解像度に切り替わった後は、動いている部分の細かな絵柄は再生されない。したがって、通常解像度に戻ったときは詳細情報を送らなければいけないが、これには多量の情報を要する。結果として、何枚もの画面がコマ落としされることになる。

これを避けるために、縮小から通常に切り替わった直後の数画面の符号化では、DCT(離散コサイン変換)演算で得られる係数のうちの一部を無視する。具体的には、切り替え後最初の通常解像度画面では、周波数の低い側の 4×4 個の係数、つぎの画面では 5×5 個、3番目は 6×6 個、4番目は 7×7 個を符号化対象とし、これらよりも高い周波数に相当する係数は強制的に切り捨てて。

● 2 : 1間引き

縮小解像度が選ばれた場合に機能するブロックで、変化部分の解像度を縦横それぞれ1/2にする。本方式では、フィルタ演算を簡易化するために、処理をブロックに閉じたものとした。具体的には、隣接する縦横2画素、計4画

素の平均値を求めて縮小画面の1画素の値としている。

● 1:2拡大

間引き同様に、縮小解像度の場合に逆DCT演算後に解像度を元に戻す働きをする。各ブロックが独立に処理できるようにするため、演算はブロックに閉じたものとした。

シミュレーション

● レート制御

低遅延の条件下で考案方式の有効性を評価するために、TMN8⁽²⁾で規定されているレート制御方式を用いた。各画面の発生情報量が設定値になるような制御である。ただし、一つの画面から大量の情報が発生した場合には、バッファ占有量を適度な値に保つため、コマ落としが起きることもある。本シミュレーションでは、フレームレートを10とした。

● 評価画像およびビットレート

独自に作成した2種類の画像“Bowling”および“Pamphlet”を評価画像として用いた。これらはいずれも、動きの大きな前景と、静止でかつ細かな絵柄がある背景を含んでいる。各画像の特徴と符号化ビットレートは、以下のようになっている。

・Bowling(48 kbps)

人物が登場し、お辞儀をして退場する{図-3(a)}。

・Pamphlet(48 kbps)

人物が登場し、パンフレットを取るために一度画面から消え、再度現れる{図-3(b)}。

● シミュレーション結果と評価

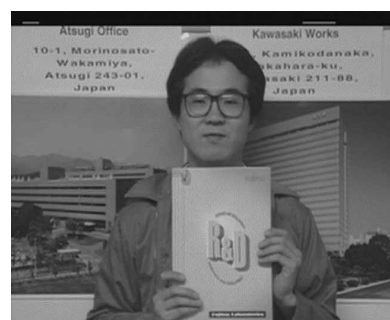
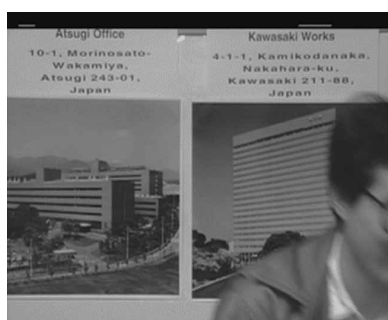
図-4(a)と図-5(a)に、各画像での画面あたりの情報発生量(各図上段の曲線)、選択された解像度(各図下段の曲線)、および画質{図-4(b)、図-5(b)}を示す。

図-4では、背景だけの画面に人物が入ってくるとき、お辞儀をするとき、画面から退場するときの3回の大きな動きがある。従来方式であるH.263では、細線の変化から分かるように、平均の3倍を超える情報が発生する。したがって、直接グラフには表示していないが、画面の枚数も減っている。また画質も下がっている。これに対して提案方式では、情報量はよく制御できており、動きに影響されることがほとんどない。これにより、フレームレートをほぼ一定に保つことができる。これは、伝送できる画面数が多いというだけでなく、遅延時間にばらつきが生じないために動きが滑らかであるという点でも効果がある。また解像度切り替えの状況を見ると、H.263方式では規定の情報量に抑えられない部分、すなわち大きな動きがある部分で縮小解像度が選択されていることがわかる。

図-5も同様に、画面に入ってくるとき、パンフレット

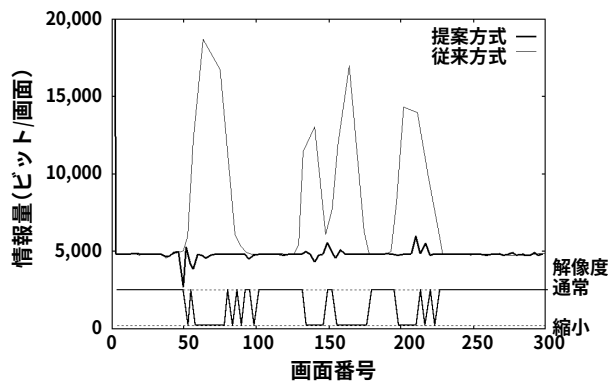


(a) 画像“Bowling”の内容

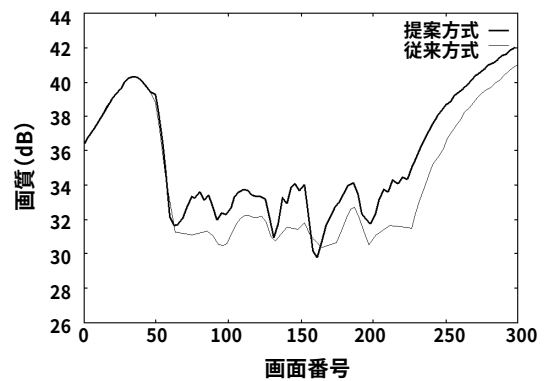


(b) 画像“Pamphlet”の内容

図3 画像“Bowling”および“Pamphlet”の内容
Fig.3-Sample pictures in “Bowling” and “Pamphlet”.

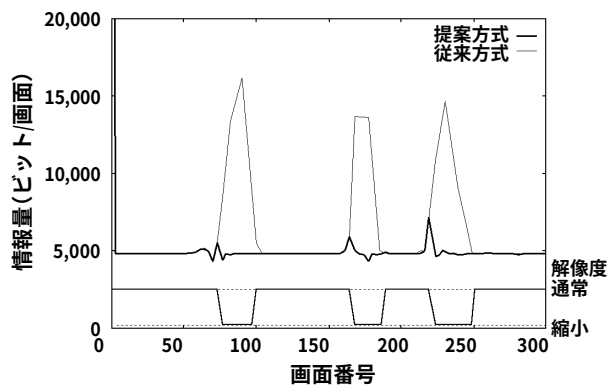


(a) 情報発生量と解像度

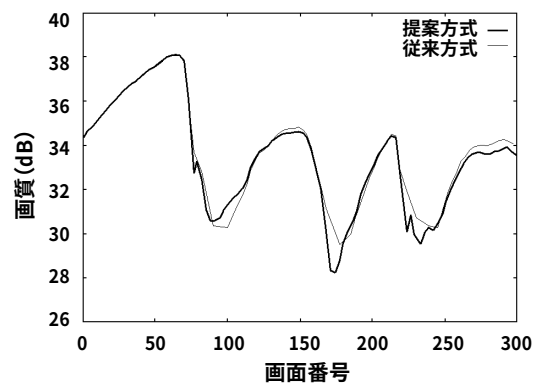


(b) 画質

図4 “Bowing”のシミュレーション結果
Fig.4-Simulation results of “Bowing”.



(a) 情報発生量と解像度



(b) 画質

図5 “Pamphlet”のシミュレーション結果
Fig.5-Simulation results of “Pamphlet”.

を取るために画面から外れるとき、持って戻ってくるときの3回の大きな動きに対応して、H.263では情報発生量にピークが生じている。この画像に対しても、考案方式では解像度切り替えが適切に行われて情報量の均一化がうまくいっている。

画質の点で言及すべきことは、動きが大きい場合に本方式で選択される縮小解像度は、本来は不利であるということである。解像度を下げ、画面数をかせぐ代わりに画質が下がるというのが、従来の符号化方式の一般的な特性であった。しかし、図4(b)、図5(b)から明らかなように、提案方式とH.263との差はほとんどない。このことと動きの滑らかさを合わせると、提案方式の方が主観品質でH.263に勝るといえる。

む す び

少ない情報量で動画を伝送する場合に品質向上を達成できる方式の一つとして、フレーム間予測方式での送

信情報である予測誤差の解像度を適応的に切り替える方式を考案した。動きが大きな画面に対して、予測誤差情報の解像度を下げて符号化する点、解像度が高い方に切り替わった直後の符号化係数に制限を設けて急激な発生情報量の変化を抑えるようにした点、解像度変換に必要な拡大・縮小フィルタ処理を簡易化した点が特徴である。

コンピュータシミュレーションにより、既存標準方式であるH.263と比べて、動きの滑らかさを保つことができ、かつ各画面の品質も同等以上であることから、主観的な品質を向上させられる方式であることを確認した。

参考文献

- (1) 伊藤ほか：低レート向け動画像符号化のための予測誤差解像度動的選択方式. 情報処理学会研究報告, 98-AVM-22, pp.25-30, 1998.
- (2) “Video Codec Test Model, Near-Term, Version 8(TMN8) Release 0”. Q15a59, ITU-T, Q15/WP3/SG16.