



Title	動き量に適応した動き補償による動画像符号化
Author(s)	顧, 毓強; 金丸, 利文; 小川, 吉彦 他
Citation	北海道大學工學部研究報告, 174, 31-39
Issue Date	1995-07-28
Doc URL	https://hdl.handle.net/2115/42452
Type	departmental bulletin paper
File Information	174_31-40.pdf



動き量に適応した動き補償による動画像符号化

顧 毓 強 金 丸 利 文 小 川 吉 彦
白 川 智 昭 北 島 秀 夫 三 木 信 弘

(平成7年3月27日受理)

An Adaptive Algorithm for Motion-Compensated Video Coding

Yuqiang GU, Toshifumi KANAMARU, Yoshihiko OGAWA
Tomoaki SHIRAKAWA, Hideo KITAJIMA, and Nobuhiro MIKI

(Received March 27, 1995)

Abstract

This paper presents an adaptive algorithm for motion-compensated video coding. The algorithm can be used for video-conferencing and broadcast television. It saves computation by conducting a variable-stage motion search according to the motion. The new algorithm is compared with a conventional algorithm (Method 2) and it is shown to attain approximately 50 percent saving in computing the motion vector. It is also shown that the algorithm can control the computation for searching motion vectors and can control the MSE for motion compensation. The performance of this algorithm is similar to the conventional algorithm (Method 2) with DCT coding when the motion is big.

1. はじめに

近年、テレビ会議、テレビ電話等の動画像通信に対する要求が高まっている。動画像の符号化においては、テレビ会議で6 Mb/s~384Kb/s、テレビ電話で128Kb/s~64Kb/sと低い符号化レートが要求される。しかしながら、動画像信号をそのままディジタル化すると膨大なデータ量になるので、大容量の伝送路が必要になり、経済的に問題になる。そこで、データ伝送や伝送路の経済化を図るため、高能率符号化によるデータ圧縮が要求され、幅広く研究がなされている。

現在、動画像の符号化で主流となっている技術は、動き補償フレーム間予測と変換符号化の一種である離散コサイン変換(DCT)符号化である。これはテレビ会議、テレビ電話信号の符号化アルゴリズムとしてITU-Tで検討されている。

図1に動画像符号化の大まかな流れを示す。この中で動き補償技術としては、いくつかの方法が提案されており、3段階の動きベクトルの探索法³⁾や4分木階層を用いた動きベクトルの検出法¹⁾等がある。4分木階層を用いた方法は動き量検出の精度を上げるために用いられる。本論文では、4分木階層を用いて3つの方法で動きベクトルを検出し、符号化特性を検討する。さらに、動き量に適応した動き補償による動画像符号化を提案する。

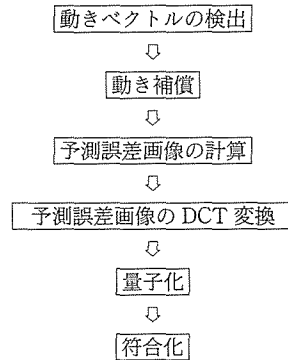


図1 動画像符号化の流れ

2. 従来の動きベクトルの検出法

2.1 4分木階層画像による動き探索

動きベクトルの検出法として広く用いられているブロックマッチング法は、その探索範囲を広げればより大きな動きを検出できるようになるが、探索に要する計算量が増大するという欠点を有する。そこで、階層画像を用いて、動きを検出する方法が提案された。これは、現フレームの画像から解像度の異なるいくつかの画像を階層的に作成し、それぞれに対してブロックマッチング法を用いることにより少ない計算量で大きな動きにまで対応するものである。ここでは、ある階層の画像の4画素の平均を1つ上の階層の1画素とする4分木階層画像を用いる。実際に動きの検出をする場合に各階層で動きを探索する範囲をどのようにすれば良いかが問題となる。ここでは、次の2つの方法を考える。

2.2 従来方式(1)

3つの階層を用い、ブロックサイズを 8×8 pelsとする。探索は各階層で 3×3 pelsの全探索を行う。 $h+1$ 階層の検出結果の2倍を h 階層における初期変位とする。階層は $h=0, 1, 2$ の3階層とし、高精度な検出を目指すため近傍ブロックの参照を行う。この方法は計算量は少ないという利点があるが、探索精度は低い。

2.3 従来方式(2)

各階層での探索範囲を大きくすれば、動きの検出精度は向上するが、探索範囲を変えて実験を行ったところ、最下位階層で3段階探索、上の2つの階層で2段階探索を行えば、精度として十分なものが得られることがわかった。そこで、計算量は増加するが、この手法で動き探索を行う方法を従来方式(2)とする。

3. 動き量を考慮した適応的な探索法

上述の従来方式(1)、(2)はそれぞれ探索の高速性、精度に重点を置いた検出法である。実際の動画像では、1枚の画像の中で動きの大きな所と小さな所が存在する。そこで、動き検出対象ブロックの動きの大きさによって従来方式(1)と(2)を切り替えられれば効率的である。動きの大きさを

判定する指標としては、動き検出を行う際に求められるマッチング評価関数（ここでは、絶対誤差を用いる）を利用するのが妥当であると思われる。以下の式によりマッチング評価関数を求める。

$$MAE(x, y) = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \{|A_k(i, j) - A_{k-1}(i+x, j+y)|\} \quad (1)$$

ここで、ブロックサイズを $m \times n$ とし、 $A_k(i, j)$ は現フレーム中の処理ブロックにおける座標 (i, j) の画素値、 $A_{k-1}(i+x, j+y)$ は前フレームにおける座標 $(i+x, j+y)$ の画素値である。

次に動きの大きさを判定するしきい値の決定法を検討する。まず、各階層において1段階の動きベクトル検出（下の2つの階層では4つの近傍ブロックの参照によって探索精度を上げている）をブロックマッチング法で行い、その際に計算されるマッチング評価関数の値を保持する。各フレームの階層毎にマッチング評価関数の最大値を求めてその平均を B （階層毎に B_2, B_1, B_0 ）とする。この B を基準としてしきい値を決定することにする。しきい値が小さくなるほど動きが大きいと判定されるブロック数が増加するため、動きベクトル検出に必要な計算量は増加する。ここでは、比較的動きの小さな画像 Walk から B の値を計算した。しきい値を決定するために、しきい値を B の値を中心として変化させて、動きベクトルの精度と必要な計算量を実験的に求める。

実験には、動画像 Walk の Y 信号の第1～第14フレーム、Football の第1～第9フレーム、Flower Garden, Cheer Leader の第1～第15フレームの4種類の画像を用いて先に説明した3つの階層の判定条件を B の定数倍としてこの倍率で変化させて、MSE とマッチング評価関数の計算回数を調べた。MSE は次の(2)式により求められる。以下のグラフで示した MSE とマッチング評価関数の計算回数は画像毎の平均値である。図2より、MSE とマッチング評価関数の計算回数の平均値の大きさを合わせて考えると、判定条件 B （階層毎に B_2, B_1, B_0 ）とする場合が最適の判定条件であることを確認した。

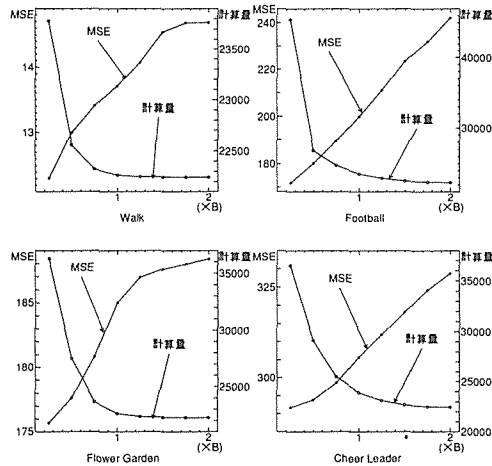


図2 判定条件 B の変化による MSE と計算量への影響

次にここで提案する探索方法を図3で説明する。まず、上の階層の各ブロックについて動きを探索する。求めたマッチング評価関数の値 M を判定条件 B_2 と比較し、大きい場合は2段階探索する。小さい場合はすぐ探索を終了する。上の階層の探索が全部終わったら、次に一つ下の階層を処理する。同様の処理を一番下の階層まで繰り返す。ただし、一番下の階層では状況によって3段階探索をすることがある。

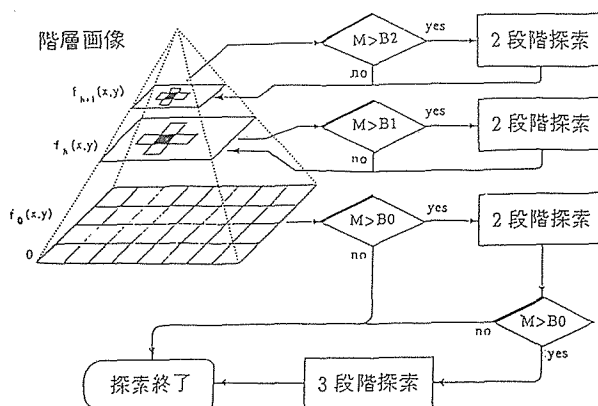


図3 提案する探索方法の説明

4. 実験方法と結果

実験には、動画 Walk の Y 信号の第1～第14フレーム, Football の第1～第9フレーム, Flower Garden, Cheer Leader の第1～第15フレームを用いた。それらの画像はいずれも 352×240 の画素の画面サイズである。以下3つの方法で動き探索を行ったときの MSE とマッチング評価関数の計算回数調べる。

1. 従来方式(1)(Method 1)
2. 従来方式(2)(Method 2)
3. 提案方式(Proposed Method)

ただし、以下の式により MSE を求める。

$$MSE = \frac{1}{352 \times 240} \sum_{i=0}^{351} \sum_{j=0}^{239} \{F(i, j) - G(i, j)\}^2 \quad (2)$$

ここで、現フレームを $F(i, j)$ 、現フレームの予測画像(動き補償をした画像)を $G(i, j)$ としている。

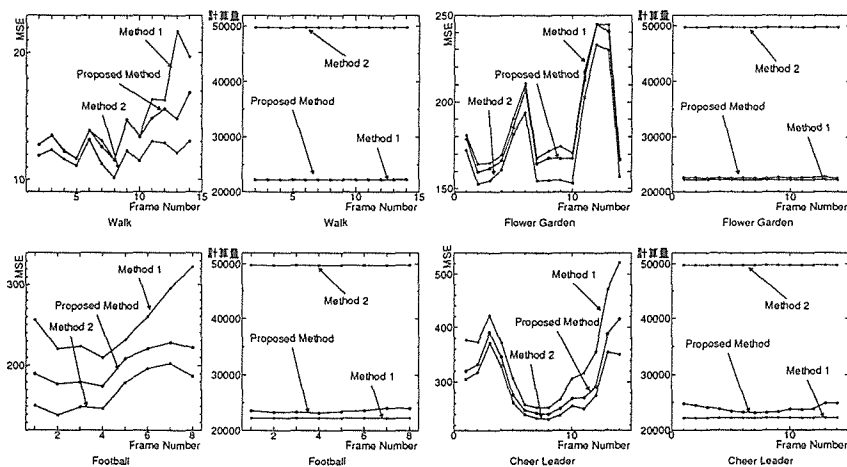


図4 MSE とマッチング評価関数の計算回数

4種類の画像を用いてフレーム毎にMSEとマッチング評価関数の計算回数を調べた結果を図4に示す。従来方式(1)は計算量が少ないという有利点があるが、MSEの値は大きい。従来方式(2)では計算量が多いという欠点があるが、MSEの値は一番小さい。提案方式は計算量がそれほど多くないが、MSEの値はほとんど従来方式(1)と従来方式(2)の中間に位置する。従って、従来方式よりも提案方式の方が効率の良い動きベクトル検出ができ、提案方式の有効性を示すことができた。

5. 動きベクトルの可変長符号化

動画像Footballの第1～第9フレーム、Cheer Leaderの第1～第15フレームのY信号を用いて以下3つの方法（ブロックサイズ 8×8 pels）で動きベクトルの伝送に必要なビットレートを調べた。動きベクトルには動画像通信用の映像符号化標準であるITU-T勧告H.261に基づき直前の動きベクトルとの差分に対して可変長符号化を行なった。結果を図5に示す。ここで、Method 1は従来方式(1)によるもの、Method 2は従来方式(2)によるものである。

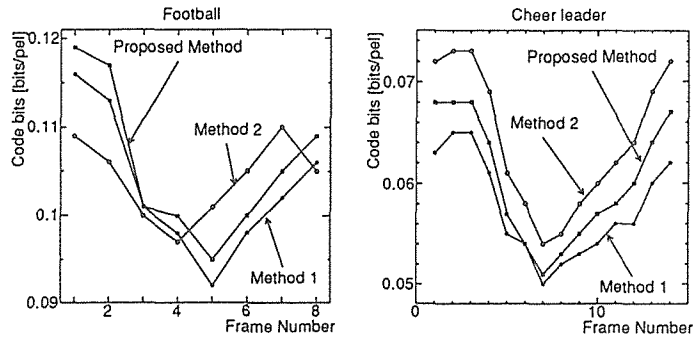


図5 ビットレート

図5より、3つの方法で求められた動きベクトルの伝送に必要なビットレートは多少の差があるが、平均値としてはほとんど同じである。提案法は必要な演算数が従来方式(2)よりかなり減少したにもかかわらず、動きベクトルの伝送に必要なビットレートはほとんど増えていないことにより、提案方式の有効性が確認できた。

6. 予測誤差信号の符号化

静止画像の符号化はDCT符号化が一般的であるが、動画像の符号化に対しても、静止画像の符号化技術が利用されている。今回は動画像Football, Cheer Leader, Walk, Flower Gardenのそれぞれの8フレーム分のY信号を用い、以下の3つの方法で符号化特性を調べた。ここで、求められたSNRはそれぞれの画像の平均値である。

1. 従来方式(1)(Method 1)+DCT符号化
2. 従来方式(2)(Method 2)+DCT符号化
3. 提案方式(Proposed Method)+DCT符号化

ただし、DCT符号化はブロックサイズを 8×8 pelsとし、ラプラスモデル(Midriser型量子化器)非線形量子化を行なった。標準偏差は4bitで線形量子化をした。次の式によりMSEを求める。

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \{A(i, j) - B(i, j)\}^2 \quad (3)$$

ここで、画面サイズを $m \times n$ 、原画像を $A(i, j)$ 、DCT 変換符号化後の復元した画像を $B(i, j)$ としている。次の式により SNR を求める。

$$SNR = 10 \log_{10} \frac{255^2}{MSE} \quad (4)$$

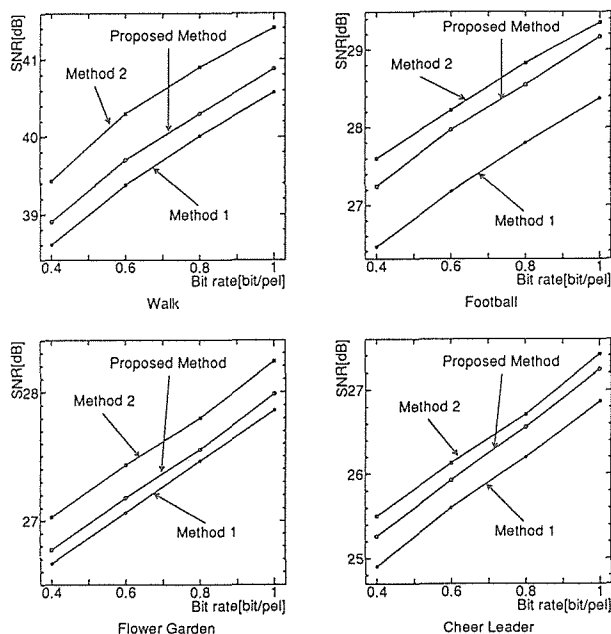


図6 符号化特性

図6より、従来方式(1)は計算量は少ないが、SNRの値は一番低いのに対し、従来方式(2)では計算量は多いが、SNRの値は一番高いことがわかる。提案方式では計算量はそれほど多くないが、SNRで従来方式(2)と比べると、WalkとFlower Gardenの画像の場合は差が多いが、FootballとCheer Leaderの画像の場合は差が少ない。提案方式と従来方式(1)を比べると、計算量はそれほど多くないが、SNRの値では、Footballの画像の場合、0.77dB程度、Cheer Leaderの画像の場合、0.35dB程度、Walkの画像の場合、0.31dB程度、Flower Gardenの画像の場合、0.11dB程度、それぞれ向上した。4種類の画像を用いて実験より今回提案した方法はSNRが向上した。特に動きの大きい画像Footballに対してよい結果が得られた。

次に動画Footballの第8～第9フレームのY信号を用いた。3つの動きベクトルの検出方法で動き補償した画像と原画像との差を調べた。差をはっきりと見るために、特に差が大きい部分、横方向第1～第186画素、縦方向第1～第160画素からなる矩形領域で比較する。

図7より、提案方式と従来方式(2)の動き補償した画像の中の左の人の胸のところは原画像とほとんど変わらない。しかし、従来方式(1)の方はその部分のひずみがはっきりとみえる。

次に予測誤差信号を3つの方式でDCT符号化をした。図7と同じように差が大きい部分、横方向第147～第333画素、縦方向第40～第200画素からなる矩形領域で比較する。



原画像



原画像



proposed method SNR=24.670 (dB)



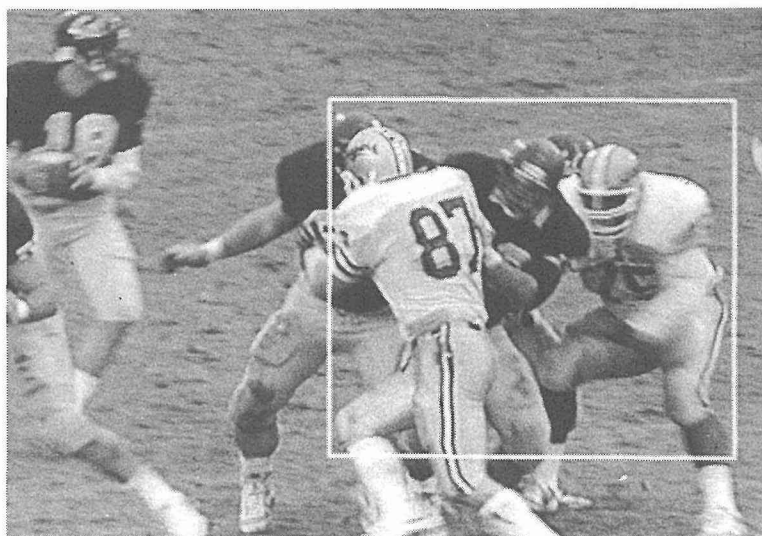
Method(2) SNR = 25.412 (dB)



method(1) SNR = 23.031 (dB)

図7 原画像と動き補償した画像との比較

M1=Method1; M2=Method 2; PM=Proposed Method



原画像 (8.00dit/pel)



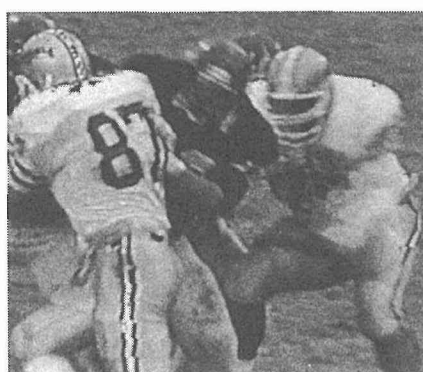
原画像 (8.00bit/pel)



pm SNR=28,799 (dB) ; (1.11bit/pel)



M₂ SNR = 28.083 (dB) ; (1.11bit/pel)



M₁ SNR = 27.732 (dB) ; (1.11bit/pel)

図8 原画像と復元した画像との比較

図8より、提案方式と従来方式(2)の復元した画像の中の右の人の腹部は、原画像と比較すると、ひずみは多くないが、従来方式(1)の方はその部分のひずみが多くなった。右の人の肩のところではひずみはほとんど同じであるが、従来方式(1)の方はその部分もひずみがある。提案方式と従来方式(1)と原画像を比較すると、左の人の頭の下のところは、ひずみはほとんど同じであるが、従来方式(2)の方はその部分のひずみが少ない。提案方式の計算量は従来方式(2)と比べると半分以上に減少し、求めた SNR は従来方式(2)とほとんど同じである。これにより提案方式の有効性がわかった。

7. ま と め

本論文では、画像の局所的な動きの大きさにより探索範囲を適応的に切り替える動き検出法を提案した。この手法によれば動きの小さな部分で無駄な探索を行わないことにより計算量を抑えながら高精度な動き探索が可能となる。また、動画像符号化の観点からも、提案法は特に従来多量の計算が必要であった動きの大きな画像に対して有効であることを示すことができた。

参 考 文 献

- 1) 富永英義, 小松尚久, 宮下壮史, 花村剛: “階層画素情報を用いた動画像における動き量検出方式”, 信学論(D-II), J72-D-II, No.3, pp.395-403 (平1-03).
- 2) 山口博久: “ラプラス分布信号の最適量子化”, 信学論(B), J67-B, No.2, pp.141-146(昭59-02).
- 3) 古閑敏夫, 平野昭, 大木淳一, 飯沼一元: “会議テレビ信号の動き補償フレーム間符号化”, 信学技報, IE81-54 (昭56).
- 4) 金丸利文, 北島秀夫, 白川智昭, 小川吉彦: “ブロックマッチング法による画素毎の動きベクトル検出法”, 信学技報, IE92-91, (平4-12).
- 5) 顧毓強, 金丸利文, 小川吉彦, 白川智昭, 北島秀夫, 三木信弘: “動画像の動き量に適応した動きベクトル検出法”, 平成6年度電気関係学会北海道支部連合大会, 334 (平6-10).