

## 領域分割された動画像におけるオブジェクト単位での圧縮方法 A Compressed method using object units for divided dynamic picture images

進 隆治† 森 秀樹† 上原 稔†

Ryuji Shin Hideki Mori Minoru Uehara  
東洋大学 大学院 工学研究科 情報システム専攻

### 1 はじめに

動画には MPEG などの標準方式と呼ばれる技術が使われている。この標準方式はあらゆる状況に対応するため高い精度で計算が行われている。そのため場合によっては過剰な処理を行うことがある。特にネットワークを利用する映像配信においては無駄な処理を少なくすることが望ましい。ブロードバンドの普及に伴い、さまざまな形で動画を楽しむ環境が現れてきた。例えば動画投稿サイトの登場により、今まで受け取り手だった人たちが作品を投稿し自己を表現することが可能となった。これらの動画には一般のテレビ放送と比べ、高負荷を必要としない映像が多く存在している。本論文では、高負荷を必要としない映像、特に合成映像に焦点を当て調査と検討を行った。

### 2 関連技術

**2.1 固定ビットレートと可変ビットレート**  
転送する情報量が常に同じであるものが固定ビットレート (CBR) と呼ばれ、可変ビットレート (VBR) とは、動きの多いシーン (高い転送量が必要) では高ビットレートにし、逆に動きの少ないシーン (低い転送量で十分) では低ビットレートに設定を行うものである [1]。

#### 2.2 GOP(Group of Picture)

GOP は時間軸上に並んでいるフレームをグループ化した構造。I・P・B ピクチャと呼ばれる3つのピクチャから構成される。

- ・I - ピクチャ - 現在の画像
  - ・P - ピクチャ - I ピクチャから予測
  - ・B - ピクチャ - I・P ピクチャ双方のデータから予測
- GOP は、画像を圧縮する際に前後の画像を参照し現在の画像を予測、その予測との誤差をデータとするため大きなデータの削減となる。

GOP 配列の長さは決まっておらずエンコーダーの設定によって決定される。基本設定では18フレーム、Iピクチャ1つにつきPピクチャ5個、Pの間にBピクチャ2個ずつ入る [1]。

基本的な GOP 配列は以下の通りである

「IBBPBBPBBPBBPBBPBB」

#### 2.2 動き補償予測

GOP は動き補償予測によってフレーム間の予測が行われている。動画の時間単位におけるフレーム間には高い相関がある。動き補償予測では次のフレームをあらかじめ予測しデータ差分を少なくしている。計算は差分絶対値和

(Sum of Absolute Difference : SAD) が使われる [2]。

$$SAD = |S_n(x) - S_{n-1}(X+V)|$$

B : 被予測画像領域  $S_n(x)$  :  $X = (i, j)$  における画素値  
V : 動きベクトル

### 3 合成映像の比較検討

合成映像には、合成元の映像の特徴がそれぞれ大きく異なる場合がある。例えば、静止画と動画を組み合わせた映像がその最たるものである。

そこで以下の図 1 のような合成映像を作成した。動画・静止画単体での映像とすでに一体化させてある同じ映像との圧縮までにかかった時間と情報量を比較した。サンプルには動きの多い映像と少ない映像を使用し、動きに対する計算の負荷の差がどれほどになるのか調べることにした。

(静止画映像は JPEG を 30 秒の動画 (MPEG) へ変換したものである)

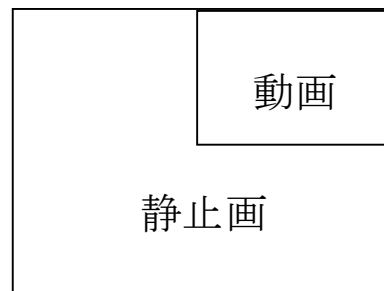


図 1 合成映像

#### 動作環境

CPU : Pentium4 - 2.8 GHz

メモリ : 1 GB

#### 圧縮設定

可変ビットレート

時間 : 30 秒

画面 720×480

表 1 圧縮にかかったエンコード時間

	動画単体	静止画単体	一体化映像
動き(多)	17(秒)	4(秒)	31(秒)
動き(少)	6(秒)	4(秒)	33(秒)

補足

静止画映像の元となった JPEG 画像の情報量は、動き(多)1.25MB 動き(少)1.34MB

静止画単体-[720×480×4分の3] 動画単体-180×120  
よって単体を合成した場合の総情報量は [動画単体]+[静止画単体]

表 2 圧縮後の情報量

	動画単体	静止画単体	一体化映像
動き(多)	15.5 (MB)	7.1(MB)	23.9(MB)
動き(少)	3.6 (MB)	5.4 (MB)	14.1(MB)

#### 4 比較と考察

エンコード時間については、一体化映像を圧縮するよりもそれぞれの映像単体で圧縮したほうが速いことが確認された。特に、動き(少)の一体化映像と単体でのエンコード時間では約三倍の時間差が現れた。動きのある部分とない部分の領域がはっきりしている場合、分離してからのほうが少ない時間で圧縮可能である。

情報量については、動き(多)では一体化映像 23.9MB、単体での総情報量は 21.1MB と大きな情報量の差はなかった。動き(少)は一体化映像 14.1MB、単体での総情報量 9.9MB と約 40 パーセントの差が現れた。

動きの少ない映像ほど、一体化映像を圧縮するよりも分離した状態でそれぞれ圧縮したほうが少ない時間で高圧縮処理を行うことができると確認された。逆に動きの多いシーンが多い場合は一体化映像、つまり分離せずに画面全体を圧縮するのが適当である。

静止画映像は、JPEG から MPEG 30 秒へと変換したところ 1.25MB から 9.5MB、1.34MB から 7.2MB と約 8 倍・5 倍ほどの情報量の差が現れた。これらの映像は動きがゼロであるため映像そのもののフレーム数を少なくすることで情報量の低減を図ることが望ましいと考えられる。

#### 5 ビットレート比較

エンコード時間・情報量のほかにビットレートについても比較を行った。前述したとおり VBR では動きの大小により、与えられるビットレートは変わってくる。一体化した場合・分離した場合のビットレートの流れについて比較を行い、最適な圧縮方法について調査した。

図 2 は上の波形が一体化映像、下の波形が動画単体でのビットレートである。

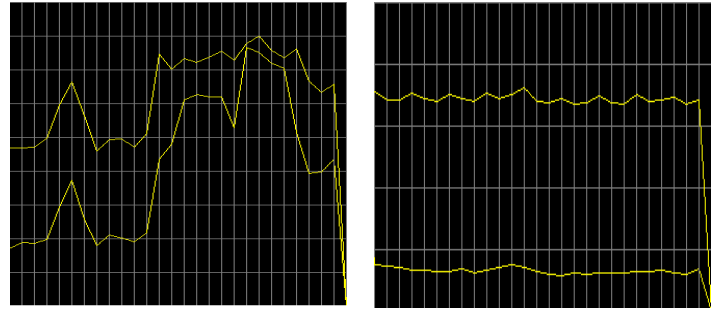


図 2 ビットレートの流れ

左：動き(多) 右：動き(少)

静止画単体でのビットレートは一定の値となり、CBR で圧縮した場合と同じ結果となった。動き(少)の場合、ビットレートの流れでは上下の変動が少なく CBR での圧縮と大差がない結果となった。動き(多)の場合、一体化動画のビットレートは動画単体でのビットレートの流れに依存する形となった。

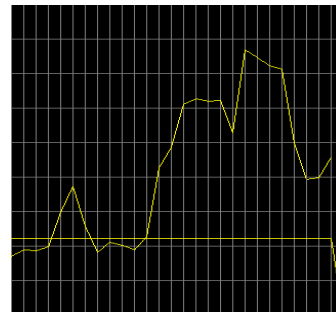


図 3 上：動画 下：静止画単体のビットレート

図 3 のように圧縮する画像領域によってビットレートの上下変動は異なってくる。「4 比較と考察」で述べたように本提案の分離圧縮法は、エンコード時間と情報量の観点からは従来の圧縮方法と同等と判断されるが、ビットレートの観点も踏まえると分離した状態で圧縮したほうが有利であると判断できる。

#### 6 まとめ

映像の種類によっては標準方式で圧縮するよりも動きの多い・少ないで分離し、それぞれ圧縮するほうが適当であることが確認できた。さらに、その判断にはエンコード時間・情報量・ビットレートなど複数の要素によって判断する必要があることを提示した。画像領域により適切なビットレート形式 ( CBR と VBR ) やフレームレートなどのパラメータは異なってくる。よって、1 つの動画ファイルに複数のファイル形式 ( MPEG + JPEG など) を組み込むことが最適であると考えられる。

#### 参考文献

- [1] 「MPEG 理論と実践」加藤 孝・鈴木 雅也 著、NTT 出版、2003 年
- [2] 「動画像の高効率符号化」小野 定康・村上 篤道・浅井 光太郎 著、オーム社、2005 年
- [3] 「H.264/AVC 教科書」大久保 榮・角野 眞也・菊池 義浩・鈴木 輝彦 著、インプレス R & D、2006 年