

符号化データによる動画解析に関する基礎検討

2D-4

熊沢 宏之 尾崎 稔

三菱電機(株)産業システム研究所

1. はじめに

マルチメディア処理では動画像を計算機で取り扱い易くすることが必須であり、そのため、動画像データに対して自動的にインデックス付けをすることが検討されている^{1)~3)}。本報告では、今後、動画像は計算機内では符号化圧縮された形で蓄積されるであろうことを考慮し、符号化データから様々な情報を抽出するための基礎検討結果について述べる。具体的には、動画像符号化の国際標準方式(H. 261, MPEGなど)で採用されている動き補償予測を想定し、予測モードから動画像のカットを自動検出する可能性について検討する。

2. 動画像符号化について—動き補償予測

動画像の符号化では、フレーム間予測により時間的冗長性を除去した信号を、直交変換符号化(離散コサイン変換: DCT)により、更に空間的冗長性を除去するというハイブリッド方式が主流になりつつある。そして、フレーム間予測は、単にフレーム間の差分を取るだけではなく、フレーム間でブロック(一般に、 16×16 画素でマクロブロックと呼ばれる)単位に動きベクトルを検出し、この動きを考慮した差分を変換符号化の対象にするという動き補償予測が用いられる。そして、フレームの性質やビットレートの制約などに応じてフレーム間予測を用いるフレーム間符号化と、直交変換符号化だけを用いるフレーム内符号化がマクロブロック単位に適応的に切り換えられる。

このマクロブロック単位の予測モード、即ち、フレーム間符号化/フレーム内符号化のいずれが用いられるかということは動画像の時間軸方向の情報を反映しており、この情報から動画像

の解析が可能になると考えられる。更に、この予測モードは符号化データから抽出できるため、極めて少ない計算量で動画像に関する様々な情報を得られるものと期待される。

3. 予測モードによるカット検出

この予測モードを用いて動画像のカットを検出することを試みる。マクロブロック単位の予測モードの切り換え手法として様々な方法があるが、ここではビットレートなどの制約はなく、予測誤差のみにより予測モードが決定されるものと仮定する。

まず、マクロブロック単位の予測誤差Dは、現フレームを $I(x,y)$ 、前フレームの復号結果を $I'(x,y)$ で表わすと、次式で与えられる。

$$D = \min_{d_x, d_y} \sum_{(x,y) \in \text{マクロブロック}} |I(x,y) - I'(x-d_x, y-d_y)|^2 / 256 \quad (1)$$

ここで、 \min は動きベクトルの探索範囲内で予測誤差が最小値をとる d_x, d_y を見つけることを意味し、この (d_x, d_y) が動きベクトルに相当する。また、動きベクトルの探索範囲は、H. 261の場合、最大で ± 15 画素であるが、この値以下であっても構わない。

このようにして求められる予測誤差Dを、以下のようにしきい値処理を行なうことにより予測モードが決定される。

$$D \leq \text{しきい値} \text{ であれば フレーム間符号化} \quad (2)$$

$$D > \text{しきい値} \text{ であれば フレーム内符号化}$$

そして、1フレーム内のマクロブロックの総数をT、フレーム間符号化されるマクロブロック数をFとすると、 F/T がカット変化の目安となる。即ち、同一カット内では似かよったフレームが連続するためFの割合が大きく、一方、カットの変化点ではFの割合が減少すると考えられる。

4. シミュレーション

マクロブロック単位の予測モードによるカッ

ト検出方式を実際の動画像を用いて検証する。但し、ここでは、実際に動画像符号化データを用いるのではなく、

- ① 動画像の各フレームをマクロブロック単位に分割
- ② このマクロブロックが、連続するフレーム間でどのように動くかを検出（動きベクトルの検出）
- ③ マクロブロック毎に予測誤差を計測し、しきい値により予測モードを決定

という方法を取る。従って、実際の動画像符号化では、動きベクトルを求める際に、式(1)における前フレームとして実際に符号化されたデータの復号結果を用いているのに対して、今回のシミュレーションでは、原画像のデータをそのまま用いている。しかし、予測モードによるカット検出の可能性の検証という目的は、今回のシミュレーションで十分に達成可能と考える。

図1に25秒間の動画像に対して、各フレーム毎にフレーム間符号化が行なわれると判定されたマクロブロック数の割合 (F/T) を示す。この実験では、NTSCの輝度信号、片フィールドから成る 256×256 画素、30フレーム/秒の動画像を作成し、これらに対して、式(1)の予測誤差を求め、式(2)の条件判定でしきい値 = 50 としてマクロブロック単位に予測モードを調べた。同図 (a) は動きベクトルの探索範囲が ± 7 画素、(b) は ± 0 画素（動き

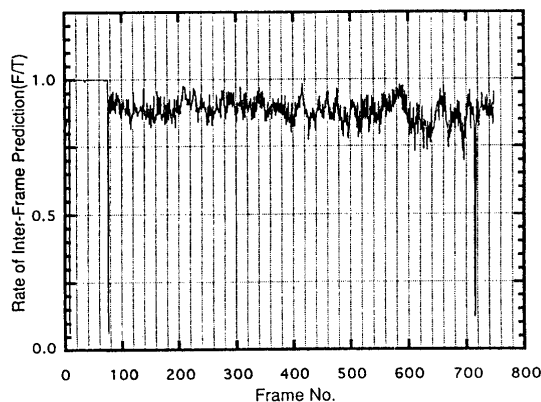
補償を行なわない場合に相当）の場合である。この動画像には、フレーム番号10、80、720付近にカットの変化があり、図1(a)、(b)共にこの点で予測モードが大きく変化しており、 F/T のしきい値を適切に選択することにより、カットの検出が可能であることを示している。また、(b)は(a)に比べて、同一カット内でも予測モードの変動が大きく、わずかな動きに対して過剰に反応している。この結果は動き補償により安定したカット検出が可能になることを示唆するものである。

5. むすび

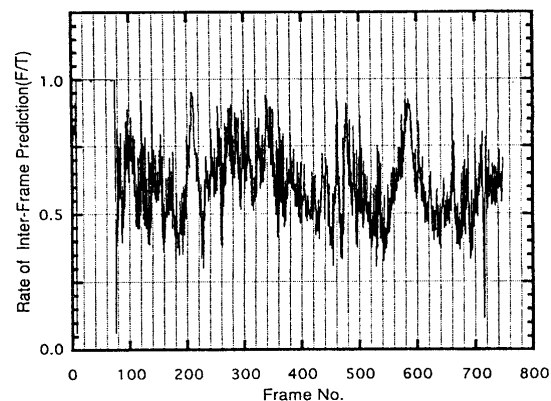
本報告では、動画像符号化データ、特に動き補償の予測モードを調べることにより、動画像のカットを検出する可能性について検討した。そして、実際の動画像データに対するシミュレーションにより、本方式の有効性を確認した。

参考文献

- 1) 長坂ほか、"ビデオ作品の場面変わりの自動判別法"、情処学40全大、pp.642-643、1990。
- 2) 上田ほか、"認識技術を応用した対話型映像編集方式の提案"、信学論Vol.J75-D-II, No.2, pp.216-225, 1992。
- 3) 大辻ほか、"輝度情報を使った動画ブラウジング"、信学技法IE90-103, 1990。



(a) 動き補償あり (± 7 画素)



(b) 動き補償なし

図1. 予測モードの変化