

5 N-4 双方向予測マクロブロックの率を利用した MPEG ビデオカット検出法の実装・評価・改良*

山下賢二, 金子邦彦, 牧之内顕文

九州大学大学院システム情報科学研究科知能システム学専攻†

1 まえがき

大量のビデオデータからの検索を行えるようにするには、1本のビデオをある単位に分割し、単位毎に実体や動きの情報を保持すると効果的である。本論文では、従来からビデオの基本単位として用いられてきた“ショットの境界（ショット）”について、MPEG ビデオのカットの自動検出法と、その処理速度・検出率の評価を報告する。我々と同様のアイデアは、我々と独立にいくつかの研究 [S96] が行なわれてきたが、本発表では実装と評価に主眼をおく。

2 MPEG ビデオの構造

MPEG では連続するピクチャ間の高い相関性を利用することで、高い圧縮率を可能としている。MPEG のピクチャには符号化の種類に応じて I, P, B の 3 種があり、一般に IBBPBBPBBI のように並びが決まっている。これら I, P, B ピクチャは、以下のように符号化の方式が異なる。

- 他のピクチャの情報を使用せず、独立して符号化される I ピクチャ（イントラピクチャ）
- 前の I ピクチャまたは P ピクチャを参照し、符号化される P ピクチャ（予測符合化ピクチャ）
- 前後双方の I ピクチャまたは P ピクチャを参照し、符号化される B ピクチャ（双方向予測符合化ピクチャ）

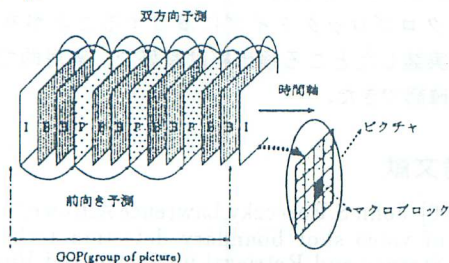


図1: MPEG の3つのピクチャ

一方、MPEG の各ピクチャは、格子状のマクロブロックから構成されている。マクロブロックには以下

*A MPEG video shot boundary detection algorithm using macroblock rate of both-direction prediction and its evaluation

†Kenji YAMASHITA, Kunihiko KANEKO, Akifumi MAKINOCHI (Graduate School of Information Science and Electrical Engineering, Department of Intelligent Systems, Kyushu University 6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-81, JAPAN)

の5種がある。符号化において、各マクロブロックごとに圧縮率の最適なものが選択される。

1. イントラブロック
(動き予測を用いないマクロブロック)
2. 前向き予測ブロック
3. 後向き予測ブロック
4. 双方向予測ブロック
5. スキップブロック
(情報を必要としないマクロブロック)

3 カット検出

当初我々は、各フレームのデータサイズの変化を利用してカット検出を試みたが、MPEG の I ピクチャは常に動き予測を用いないためデータサイズに大きな変化がなく、検出不可能な場合が起こった。

そこで、前後双方のピクチャを参照し符号化された B ピクチャに着目した。B ピクチャは、I ピクチャや P ピクチャと違い、後向き予測と双方向予測のブロックを含んでいる。カットでないところ（図2左）では、B ピクチャのマクロブロックのほとんどは、前後双方の予測ブロックである。一方、カットでは（図2右）、その部分の前後とのピクチャの相関関係がなくなり、双方向の予測を用いるブロックの使用率が減少する。すなわち、B ピクチャのマクロブロックタイプの割合がある閾値以下になるかを検出することで、カット検出が可能である（図3）。

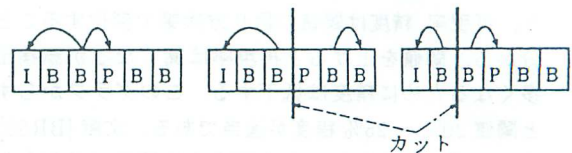


図2: カットとカットでない場合の B ピクチャ

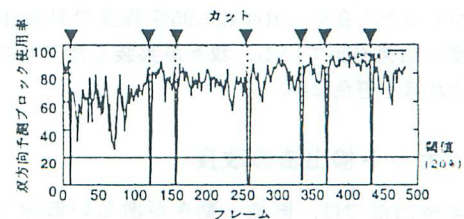


図3: B ピクチャの双方向予測使用率

4 カット検出の評価

3章に示した検出法を評価するために、5種類のビデオ（ニュース、スポーツ、教育、アニメ）に対して、カリフォルニア大パークレイ校のmpeg_encodeを用いたMPEG符号化を行い、検出率の測定を行った。測定項目は、再現率（ $\equiv \frac{\text{正しく検出したカット数}}{\text{実際に存在するカット数}}$ ）、精度（ $\equiv \frac{\text{正しく検出したカット数}}{\text{検出カット数}}$ ）の2つと、検出速度である。本実験に用いた検出プログラムはCで250行程度であり関連ツールはWWW上で公開している¹。

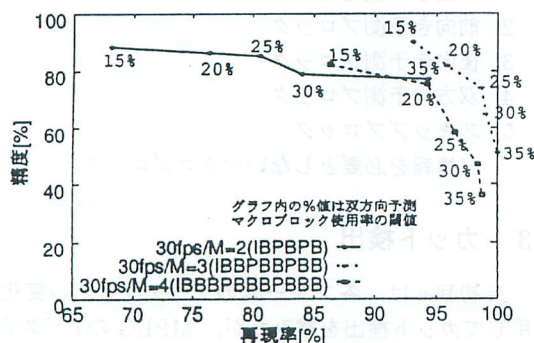


図 4: 再現率と精度

フレーム数	フレームレート	ビデオ長	平均検出速度
5000	30fps	167 秒	272 秒

表 1: 平均検出速度 (測定環境:Sun SparcStaion20)

本方式ではMPEGで符号化されたデータをそのまま利用する。したがって検出速度については、実時間に近い速度で検出が可能であり実用的な速度と言えるだろう。

検出率については、閾値を15%～35%で変化させたところ、再現率、精度も変化した。図4には5種類のビデオの検出率の平均を示している。この結果より、再現率、精度は閾値の取り方次第で変化することが分かる。閾値を上げると再現率は高くなるが誤検出が多くなるために精度は低下する。このグラフからすると閾値20%～25%程度が適当である。文献[BR86]では、ビデオの各種カット検出法の比較検討が行なわれており、RegionアルゴリズムとRunningアルゴリズムがもっとも良く、Recall=95%程度でPrecision=90%程度の結果が出ている。我々が実装したアルゴリズムもそれほど遜色ない。

5 カット検出法の改良

本検出法では、画像の動きが激しい部分で、本来カットでない部分もカットとして検出してしまう傾向がある。一方、検出洩れも少し発生している。そのため、カット検出アルゴリズムの改良が必要である。

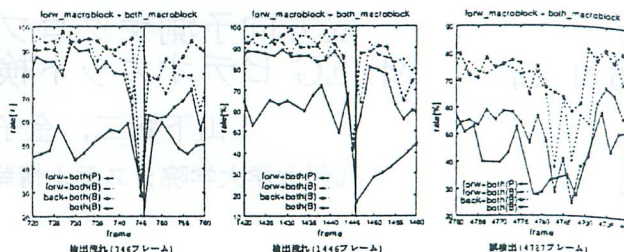


図 5: 検出洩れ、誤検出の例

図5左のグラフはスポーツ番組についての各マクロブロックの使用率を示したものであるが、双方向予測ブロックの使用率が閾値以下にならないため、検出洩れが発生している（図5左、中）では746,1446フレームがカットである）。図5右のグラフは同じスポーツ番組で誤検出が起こった部分のグラフである。ここではカメラの高速なパンと人がカメラ前を横切ったために双方向予測ブロックが閾値以下になってしまっている。

現在のアルゴリズムでは、Pピクチャの持つ前方向予測マクロブロック情報も使用していないが、図5左、中のグラフから分かる通りカットではPピクチャのマクロブロックの割合も変化するので、その利用が改良点として挙げられる。

一方、本方式はピクチャとそのピクチャが予測に使用しているピクチャとの間の時間的な距離の差を考慮していない。ピクチャ間の距離によってマクロブロックに重み付けすることで、さらに正確なフレーム間の相関関係を求めることが可能になり、これをカット検出に適用すれば有効ではないかと思われる。

6 結論

MPEGビデオに対するカット検出ではBピクチャのマクロブロックタイプに着目することが有効であり、実装したところ検出率は良好で、実用的であることが確認できた。

参考文献

- [BR96] John S.Boreczky, Lawrence A.Rowe, "Comparison of video shot boundary detection techniques", Storage and Retrieval for Image and Video Databases IV, Proc. of IS&T/SPIE 1996
- [S96] 椎尾一郎, "予測ブロック数を利用したMPEG動画画像カット検出法", 第52回情報処理学会全国大会論文集, 1996
- [OTO94] 大辻清太, 外村佳伸, 大庭有二, "突出検出フィルタを用いた映像カット点検出法", 電子情報通信学会論文誌, 1994
- [IA93] 岩成英一, 有木康雄, "DCT成分を用いた動画シーンのクラスタリング-カット検出の一手法-", 情報処理学会第47回全国大会

¹http://www.db.is.kyushu-u.ac.jp/~yamasita/mpegtool/