

映像中の特定シーンの検索方式

3E-6

長坂晃朗^{*} 宮武孝文^{*} 上田博唯^{*} 大野一広^{*}^{*}(株)日立製作所 中央研究所 ^{*}(株)日立製作所 基礎研究所

1. はじめに

近年、映画やTV番組などの映像情報を対象としたデータベースの構築が活発になっている。そして、その検索には、キーワードを用いる方法が一般的である[1][2]。しかし、映像が内包する多様な情報の全てにキーワードを付与することは事実上不可能であり、映像そのものの特徴をもとに検索を行う方法が必要になる。文献[3]では、ユーザが例示した一枚のフレーム画像上の色情報の分布を被写体の特徴量として、被写体の登場シーンを検索する方法を報告した。本稿では、シーンの色情報の時間方向の変化を特徴量として用い、全映像中の同一シーンを全て検索する方式を報告する。これにより、例えば、番組のオープニングやCMを手掛かりにした映像の検索が可能になる。また、番組やCM放送の動向を示す統計データの作成も自動化できる。

2. シーンの照合方法

シーンが同一か照合する処理の要件としては、情報量の大きな映像について実用的な時間で処理を行うために、簡単で高速なアルゴリズムであることが必須である。また、機器の特性やノイズによる映像変動にロバストであることも必要である。そこで、次の方針に基づく照合アルゴリズムを考案した。

- 1) シーンの代表フレームについてのみ照合を行う。
- 2) フレームの特徴量を10文字程の文字列で表現する。
- 3) 特徴量の照合に冗長性を持たせる。

1)により、必ずしも毎フレームについて照合処理を行う必要がなくなり、高速化が図れる。シーンはいくつかのカットの連続体であり、代表フレームと

しては、それらカットの変わり目のフレームを検出[4]して用いる。シーンの先頭は必ずカット変わりと一致するので再現性があり、代表フレームとして適している。2)は、映像の照合を文字列の照合に置き換えることにより、特徴量間の複雑な類似度計算が不要になるため、高速な照合が可能になる。また、3)は、映像変動による検出漏れを防ぐための措置である。以下では、2)と3)について詳しく述べる。

2.1 画像特徴量のコード化

コード化処理の流れを図1に示す。高速な照合を実現するため、フレームの特徴量には、RGB各色成分の平均を用いるが、フレーム全体で平均をとると特徴として不十分なので2×2に4分割し、各部分領域単位で求める。こうして求めた各成分ごとの平均値を0から100の範囲に正規化し、図1右側上に示した変換対応表に基づいて1文字の主コードに変換する。このとき、値が、対応表の範囲の中央付近ではなく、隣接した範囲との境界近傍にあたる場合には、補助コードを付加し2文字のコードとする。

2.2 コードの照合

コードの照合は、基本的に主コードが一致するかどうかで判定する。しかし、主コードは特徴量の値を機械的に範囲分けして生成しているため、境界付近であれば映像変動によって変化する可能性がある。そこで、主コードが異なる場合でも、続けて補助コードの有無を調べ、境界付近とわかれれば、その境界を成す他方の主コードと一致するかを調べて最終的な照合判定とする。

3. 実験結果とその検討

表1は、315分の連続したテレビ放送に対するCMの検索結果である。テンプレートとなるコードは、前日のテレビ放送から作成した。同じCMが複数回

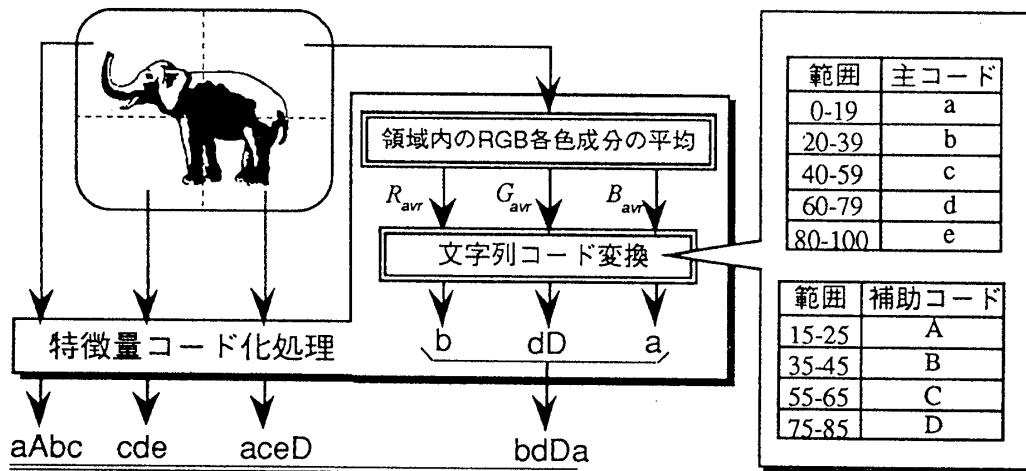


図1 フレームのコード化の概略

現れ、種類は45、総出現数は71であった。表の結果から、ほぼ実用的な検索性能が得られていることがわかる。過剰検出は無関係のシーンを検出した回数であるが、そのうち21までが全体に暗い無彩色のCMの場合が占め、時代劇やドラマなどの夜のシーンに多く過剰反応した。このようなCMの特定には、色以外の特徴を用いることが必要である。

また、アルゴリズムの特性上、カットが少ないシーンの場合、過剰検出が起きやすい。6カットから成るシーンについて、テンプレートとして用いるカ

ットの数を変化させたときの過剰検出数の変化を図2に示す。結果として、3カット以上からなるシーンには問題なく本手法が適用できることがわかった。

処理速度については、IRIS 4D310/VGX上に作成したプロトタイプでは、放送中の映像を実時間で検索できる性能が得られた。代表フレームのコードが作成された後では約1万倍速で検索でき、5時間の映像検索が2秒足らずで完了する。

4. おわりに

本稿では、映像検索の一手法として、ユーザが例示したと同じシーンを全映像中から高速に検索する技術について述べた。検索の特徴量として、画像のRGB各色成分の平均値をコードに置き換えて用いることで、映像の照合がテキスト照合のごとく単純な文字列照合として高速に行なえる。今後の課題としては、特徴量を増やし、カット数の少ないシーンの特定能力を向上させることが挙げられる。

参考文献

- [1]柴田他：“映像製作のためのシーン記述”，テレビ学技報, Vol.16, No.10, pp.19-24
- [2]安部, 外村：“状態の時間変化をキーとする動画画像検索法”，信学論, Vol.J75-D-II, No.3, pp.512-519
- [3]長坂他：“被写体に着目したビデオ映像の検索と構造視覚化”，計測自動制御学会93PG0010, pp.1-6
- [4]上田他：“認識技術を応用した対話型編集方式の提案”，信学論, Vol.J75-D-II, No.2, pp.216-225

表1 CM検索実験の結果

実験時間	315分 (3160カット)	
検索シーン種類数	45	
正認識/総出現数	71/71	
誤認識	過剰検出	25
	検出漏れ	0

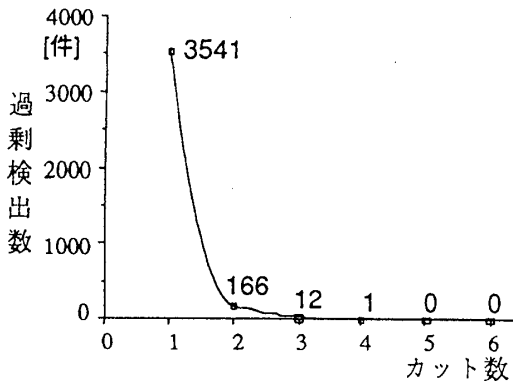


図2 カット数と過剰検出の関係