

特徴次元圧縮による長時間映像中における同一区間映像の高速検出手法

A method for fast retrieval of identical video segments in a long video stream by feature dimension compression

野田 和広†
Kazuhiro Noda目加田 慶人‡
Yoshito Mekada井手 一郎†
Ichiro Ide村瀬 洋†
Hiroshi Murase

1. はじめに

記憶装置の高性能化により、長時間の映像データを蓄積できるようになってきている。これに伴い、大量の映像を高速に検索する技術が必要とされている [1, 4, 5, 6].

従来、特定の区間映像をキーとして計算機に与え、それと同じ映像区間が、長時間の映像中のどこにあるかを探し出す映像探索手法 (図 1(a) 参照) は多数提案されている。

本手法は、これらの手法とは異なり、特定の区間映像を与えることなく、長時間映像中のどの区間とどの区間が一致しているかを映像中から全て探し出す手法である (図 1(b) 参照)。これらは同一区間映像 [3] の検出技術と呼ばれる。このような技術が実現できれば、例えば、数週間もの蓄積された放送映像から、多種の同一CMを全て検出することができる。従来のCM調査システム [2] では、キーとなる新規のCMをマニュアルで映像中から抽出し、データベース化する必要があったが、本手法により新規CMのデータベースの自動構築が可能となる。

区間映像間の照合にかかる計算量は、図 1(a) のようにキーとなる特定区間映像を長時間映像から探す場合と比較して、図 1(b) のように不特定の同一区間映像を全て検出する場合では、大幅に増加する。これは、前者では長時間映像の時間長に比例する計算量であるのに対し、後者において不特定の同一区間映像を全て検出する場合には、計算量が 2 乗オーダーで増加するためである。

本研究では、2 乗オーダー部分の処理に特徴量の次元圧縮を導入し、計算量の質的な変換を図り、全体の計算時間を大幅に短縮する手法を提案する。

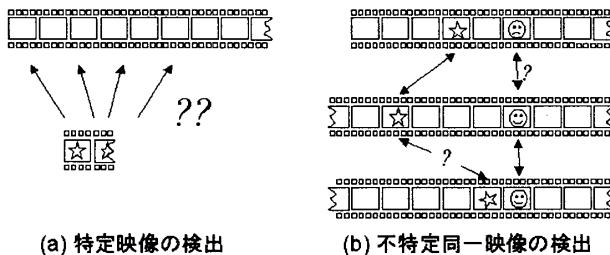


図 1: 従来の映像探索手法との照合の比較

2. 特徴次元圧縮による計算量の抑制

前述のとおり、長時間映像中の同一区間映像をすべて検出する場合の最大の問題点は、長時間映像の時間長の 2 乗オーダーで計算量が増加することである。

†名古屋大学大学院 情報科学研究科

‡京大生命システム工学部

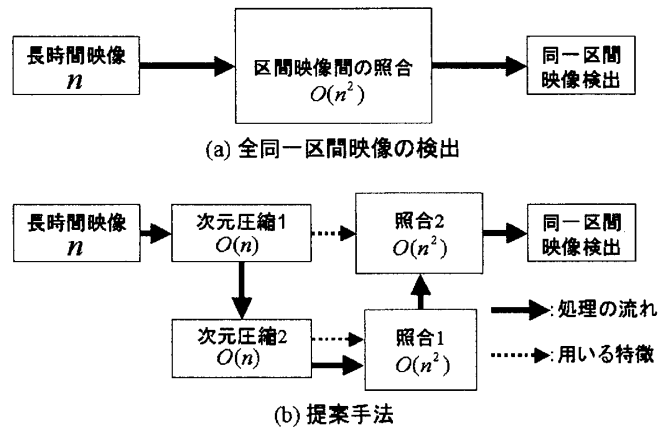


図 2: 次元圧縮による計算量の抑制

通常の映像探索ではキーとなる特定の映像を長時間映像中から探し出せばよい。キーの映像と長時間映像中の区間映像との照合を、区間を少しずつずらしながら 1 回通り行えばよいので、長時間映像の時間長に比例して計算量が増加する。一方、長時間映像中の不特定の同一区間映像を全て検出したい場合は、何の区間映像が同一で、またそれがどこに、いくつあるか、全て不定である。このような場合、長時間映像中のある区間映像をキーと考えて、その区間映像と長時間映像中の他の区間映像を少しずつずらしながら照合し、1 回通り照合したら、キーとなる区間映像を少しずつずらし、再度その区間映像と全ての区間映像との照合を繰り返すことになるので、2 乗のオーダーで計算量が増加してしまう。

このような、不特定の同一区間映像を全て検出する場合特有の問題を解決するために、区間映像の特徴次元圧縮に基づく高速検出手法を提案する。本手法は長時間映像中の各区間映像に対し、多段階の特徴次元圧縮を行う。そして低次元の特徴量を用いた照合により同一区間の候補を絞り、次の段階において高次元の特徴量で照合することにより、同一区間を検出する。次元圧縮に要する計算量は長時間映像の時間長の 1 乗のオーダーである。従って、本手法における特徴次元圧縮の意義は、長時間映像から同一区間映像を全て検出する際の時間長の 2 乗オーダーである膨大な計算量を、時間長の 1 乗オーダーの次元圧縮と 2 乗オーダーの照合に分解することで抑制する点である。つまり次の式を実現する。

$$O(n^2) \Rightarrow aO(n) + bO(n^2) \quad (a \gg b)$$

また、本手法では段階的に次元圧縮、区間映像間の照

合を行うことで、計算の効率化を図っている(図2)。

なお、特定の映像を探す、通常の映像探索に本手法を適用しても効果は少ない。なぜなら次元圧縮段階で、映像探索時に行われる処理とほぼ処理量が等価な固有ベクトルとの照合演算が入ってしまうためである。

3. 処理手順

本手法の処理手順を簡単に説明する。長時間映像のキャプチャ以前に行う準備処理と、キャプチャ以降に行う映像検出処理とに分離して述べる。

3.1 準備処理

映像のキャプチャ以前に行う処理として、事前に集めた種々の映像に対して主成分分析(固有値展開)を行う。ここでは、十分多くの映像を集めれば、一般の映像の性質を表現できると仮定している。

3.2 映像検出処理(オンライン)

映像をキャプチャした後に、まず映像のフレーム抽出を行い、次に以下で示す処理を行う。

3.2.1 特徴次元圧縮

本手法では、各区間映像に対し2段階の特徴次元圧縮を施している。まず、第1段階として映像の各フレームの特徴に対しての次元圧縮を施す。次に、第2段階として、第1段階の次元圧縮が施されたフレームを、照合する区間映像の時間分まとめたものに対して、次元圧縮を施す(図3)。なお区間映像は、事前に定めた時間長分のフレーム群を1フレームずつずらして切出し、次元圧縮を繰り返す。次元圧縮は、具体的には準備処理の固有値展開により得られた上位の主成分(固有ベクトル)を用いて、区間映像の特徴ベクトルとの内積をとり、さらに低次元の特徴を生成する。

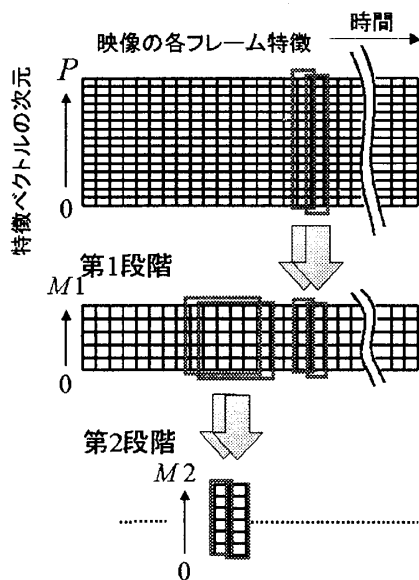


図3: 2段階の特徴次元圧縮

3.2.2 区間映像間の照合

第2段階まで次元圧縮をした映像特徴を用いて、長時間映像中の各区間映像の間で照合を行い、同一だと思われる区間映像対の候補を大まかに検出し、その候補のみに対し、第1段階の次元圧縮段階の映像特徴を用いて照合することで、より正確に同一区間映像対を検出する。照合時の類似度として区間映像の特徴間の内積の2乗を採用し、この値がしきい値より高ければ、同一区間映像対として検出する。

3.2.3 事後処理

検出した区間映像対の位置の補正のための、同一区間の区間調整処理や、検出した区間映像を同一映像グループごとにグループ分けする処理などを行う。

4. 実験と考察

次のような仕様をもつ映像に対し、本手法を含む3手法を適用し、照合時間や検出精度などを比較、検証した。

- 同一CMが1対(2区間)含まれる約20分の放送映像
- さらに手入力で同一CMを3区間挿入
- フレームサイズ320×240ピクセルを事前処理で20×15ピクセルに低解像度化

その際、下記の3つの照合手法により比較した。

1. 圧縮をしていない映像特徴で照合
2. 提案手法1: 段階的に次元圧縮せず、第1段階の圧縮段階の映像特徴のみにより照合
3. 提案手法2: 第2段階まで次元圧縮した映像特徴による大まかな照合の後、第1段階の圧縮段階の映像特徴で詳細に照合

本実験では映像特徴として、各フレームの正規化したRGB値による多次元ベクトルを用いている。また各圧縮段階での映像特徴の次元数は、非圧縮時で900×150次元、第1段階の圧縮後で10×150次元、第2段階の圧縮後で20次元である。

照合にかかった時間と検出の精度を表1,2に示す。ここで、「類似区間検出数」とは誤って同一と判定された類似区間映像の検出数を指す。よって類似区間映像の検出数が少ないほど検出精度が高いといえる。なお、本実験において検出された区間映像が同一であるか類似しているかは人間の目で判断し、同一としたのは先に述べた5区間のCMのみである。比較した3つの照合手法では、しきい値の設定により5区間全て検出できたので、検出率は全て100%であった。表2では検出率100%になるようにしきい値を設定し、そのとき誤って検出された類似区間映像の数を比較している。

また表1において照合にかかった時間の他に、次元圧縮にかかった時間と、両方の和をとった実質的な検出時

間を示す。ここでの次元圧縮は固有値展開の時間は含まず、3.2節で述べた映像検出処理中の次元圧縮を指す。

そして今回の実験結果から予測される、1週間分の映像に対して各手法を適用したときの計算時間を表3に示す。

表1: 圧縮回数の異なる3手法の検出時間

	照合時間 (秒)	次元圧縮時間 (秒)	検出時間 (秒)
(1) 非圧縮	6493.23	0	6493.23
(2) 提案手法1	70.28	3.59	73.87
(3) 提案手法2	1.72	18.90	20.62

表2: 圧縮回数の異なる3手法の検出精度

	同一区間 検出率	類似区間 検出数(個)
(1) 非圧縮	100%	11
(2) 提案手法1	100%	15
(3) 提案手法2	100%	15

表3: 1週間分の映像からの同一区間映像計算時間の予測

	予測される 計算時間(時間)
(1) 非圧縮	458,162.31
(2) 提案手法1	4,959.46
(3) 提案手法2	124.01

表1の通り、映像特徴の次元圧縮により若干の類似映像検出の増加があるものの、照合時間は大幅に減少している。一方で次元圧縮にかかる時間は増加しているが、予測される1週間分の検出時間を見ると、さらに大きく差が開き、高速化の効果が期待される。これは本手法が2乗オーダーの検出処理を、1乗オーダーの次元圧縮と2乗オーダーの照合に分解して、2乗オーダーの基本処理の量を低減しているからである。図4は、より長時間の映像に各手法を適用したときに、予測される検出時間である。図のように、適用する長時間映像の時間が長くなるほど、1乗オーダーの効果が大きくなることを期待される。

また次元圧縮により、表2のように若干類似映像の誤検出が増加したが、実際に検出した類似映像を見たところ、検出はやむを得ないと判断しうる程度の誤りであった。

しかし、3.1節で述べた固有値展開において、第1段階の固有値展開に約29分、第2段階の固有値展開に約110分と膨大な時間がかかった。ただし、提案手法を蓄積メディアに対して適用するのであれば、固有値展開は事前に一度行えば済むため、検出時の問題にはならない。

また図4のように、提案手法は1週間の映像に適用した場合でも実時間以下の時間で検出できると予想されるが、映像キャプチャ時に並行的に次元圧縮を行えば、検出処理はさらに高速になる。

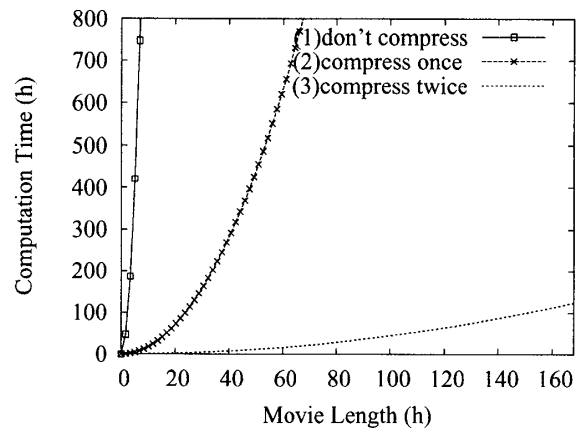


図4: 長時間映像に適用したときに予測される計算時間

5. おわりに

本稿では、長時間映像から全ての同一区間映像を高速に検出するための特徴次元圧縮に基づく手法を提案した。映像時間長の2乗オーダーで計算量が増加する検出処理を、1乗オーダーの次元圧縮と2乗オーダーの照合に分解し、2乗オーダーの部分の基本計算量を低減することで、検出を高速化できることを示した。実験と考察により、1週間程度の長時間の映像に対し適用したときに、大幅に計算量を削減できる見通しを示した。また、次元圧縮による同一区間映像の見落としではなく、若干誤って検出した類似区間映像の個数が増えたものの、照合にかかる時間を大幅に削減できた。

今後の課題として、より長時間の映像での実際的な評価、従来の映像探索手法との併用、新たな特徴次元圧縮法の検討等を考えている。

謝辞

日頃熱心に討論していただく村瀬研究室諸氏に感謝する。本研究の一部は、文部科学省・日本学術振興会科研費及び、COE補助金による。

参考文献

- [1] 村瀬洋: “画像、音、映像を高速に探索する技術 アクティブ探索法”, HCG研究会, 2002.3
- [2] 柏野邦夫, 村瀬洋: “音や映像を瞬時に探すアクティブ探索法”, NTT R&D, vol.49, no.7, pp.407-413, 2000
- [3] 山岸史典, 佐藤真一, 浜田喬, 坂内正夫: “大規模映像アーカイブにおける同一映像断片探索の高速化”, 画像ラボ, pp.15-18, 2002.8
- [4] 木村昭悟, 柏野邦夫, 黒住隆行, 村瀬洋: “区分線形写像に基づく次元削減を導入した音や映像の高速探索”, 信学論, vol.J85-D-II, No.10, pp.1552-1562, 2002.10
- [5] 木村昭悟, 柏野邦夫, 黒住隆行, 村瀬洋: “音や映像の高速探索のための動的分割に基づく特徴次元削減法”, 信学技報, PRMU2002-167, pp.127-131, 2002.12
- [6] 田村秀行編著: “コンピュータ画像処理”, オーム社, 2003