

特徴空間のマージン付き分割による繰り返し映像区間の高速検出

鈴木 秋吾[†] 高橋 友和[†] 井手 一郎^{†,††} 村瀬 洋[†]

[†] 名古屋大学大学院情報科学研究科 〒 464-8601 名古屋市千種区不老町

^{††} 国立情報学研究所 〒 101-8430 東京都千代田区一ツ橋 2-1-2

E-mail: †{shsuzuki,ttakahashi,murase}@murase.m.is.nagoya-u.ac.jp, ††ide@nii.ac.jp

あらまし 近年、ユーザの要求する映像区間を大量の映像から探索する技術に関する研究が盛んに行われている。特に放送映像においては、繰り返し放送される映像はコマーシャルや同一素材ニュース映像など特別な役割を果たすことが多く、これらを映像の意味構造解析に用いる試みが提案されている。しかし、事前に分かっていない繰り返し映像区間の検出処理には、対象とする映像長の2乗の計算量を要する。そのため、長時間映像に適用するには膨大な計算時間がかかる。本報告では繰り返し映像区間の高速検出を目的として、映像区間の照合が行われる特徴空間を分割することで、照合回数を削減する手法を提案する。空間分割では分割境界による空間の2分割を階層的に行う。加えて、各分割境界付近にマージン区間と呼ばれる保証区間を設け、両側の空間を部分的に重複させることで、分割前と同じ検出精度を実現する。実験により、提案手法による照合時間の大幅な削減を確認し、繰り返し映像区間の高速検出に対する提案手法の有効性を示した。

キーワード 空間分割, 繰り返し映像, 高速化, 精度保証

Fast Detection of Near-Duplicate Video Sequences by Division of Feature Space with Margin

Shugo SUZUKI[†], Tomokazu TAKAHASHI[†], Ichiro IDE^{†,††}, and Hiroshi MURASE[†]

[†] Graduate School of Information Science, Nagoya University
Furo-cho, Chikusa-ku, Nagoya-shi, Aichi, 464-8601 Japan

^{††} National Institute of Informatics

2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, 101-8430, Japan

E-mail: †{shsuzuki,ttakahashi,murase}@murase.m.is.nagoya-u.ac.jp, ††ide@nii.ac.jp

Abstract Recently, methods to search a video sequence from massively stored video data is being actively developed. As to broadcast video, a video sequence which is broadcasted repeatedly plays a special role, for example, a commercial or a video used in the same news topic. Detection of such duplicate video sequences, however, requires a large computational cost that is proportional to the square of whole video length. In this report, we propose a method that divides a feature space to detect duplicate video sequences rapidly. In the space division, dividing the space into two regions by a division boundary is executed hierarchically. In addition, to guarantee the same detection accuracy as that is obtained from the original space, a margin is set near the boundary. Since the experimental results showed that drastic reduction in computational time was achieved by the proposal method, the effectiveness in fast detection of duplicate video sequences was demonstrated.

Key words Feature space division, duplicate video sequence, fast detection, accuracy guarantee

1. ま え が き

長時間の映像を記録することが可能となり、効率的な映像探索技術が求められている。その一環として、特定の映像区間をキーとして、それと同一の映像区間を入力映像中から高速に探

索する技術が多く研究されている [1] [2]。一方、それとは別の取り組みとして、特定の映像区間をキーとして与えるのではなく、入力された長時間映像中に繰り返し現れる映像区間をもなく検出する技術に関する研究や、それを映像の理解に役立てる研究が盛んに行われている [3] [4] [5] [6]。この技術を用いて、

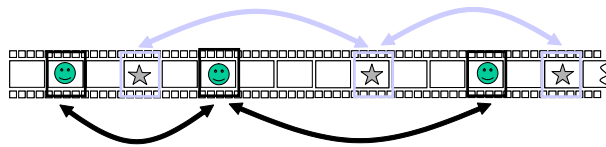


図 1 繰り返し照合

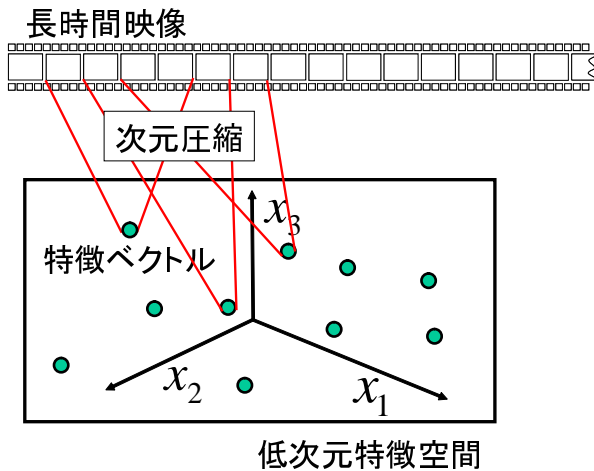


図 2 次元圧縮処理



図 3 繰り返し映像区間の例

社本ら [6] は関連のあるニューストピックの意味構造を抽出する手法を提案している．また，小川ら [5] はやはり検出された映像区間を手がかりに多言語間の同一ニュースを検出する研究を行っている．

しかし，図 1 に示すように，繰り返し映像区間を検出するには全映像中のとりうる全ての区間対を繰り返し照合しなければならないため，膨大な計算時間を要する．これに対し，野田ら [4] は，特徴次元圧縮によって映像区間を表す特徴ベクトルの次元数を減らし，繰り返し行われる各照合の計算時間を短縮している（図 2）．しかしながら，この手法では，照合回数は映像区間数の 2 乗のままであるため更なる高速化が求められる．

そこで我々は，特徴空間の分割を用いた繰り返し映像区間の高速検出手法を提案する．空間分割のアプローチは最近傍点探索問題において幅広く研究が用いられている．Arya ら [7] は特徴空間に対して，木構造を用いた空間分割を適用し，照合回数の絞込みを行い，キーとして与えた点と最も近い点の高速検索を実現している．荒井ら [8] は，上記の手法と合わせて主成分分析を用いることで，適切な分割軸を設定し，より高速な検出を行っている．

提案手法では，特徴空間の空間分割を導入することによって照合する映像区間の数を制限し，照合回数の削減を行うものである．さらに，マージン区間と呼ばれる区間を分割境界付近に設けることによって，分割の際に生じる分割境界付近の検出漏れを防ぐことができ，分割前と同じ検出精度が保証される．

2. 問題定義と従来研究

本節では，まず繰り返し映像区間検出の問題を定義した上で，従来手法である特徴次元圧縮による高速手法の概要を紹介する．

2.1 繰り返し映像区間検出の問題定義

本研究では，定量的に繰り返し映像区間を評価するため，以下の条件を満たす映像区間対を全て検出することを目的とする．

入力映像中の全映像区間を特徴ベクトルとして表現し， p, q を任意の映像区間をあらわすベクトルとする．
このとき， $\|p - q\| \leq \theta$ を満たす
 p, q を繰り返し映像区間と定義する．

単純な手法では， d 次元のベクトル間の距離を計算するのに， d に比例する計算コストが必要であり，また， N 個のベクトルの中から全てのベクトル間の距離を算出するには計算回数が N^2 となる，つまり高次元になるほど，またベクトル数が大量になるほど探索に時間が掛かる．

図 3 に本手法で検出可能な，繰り返し映像区間の例を示す．これら全てを検出することには多様なニーズがある．まず，(a),(b) にまったく同一の映像区間を挙げる．これには，コマーシャル，番組宣伝，番組のオープニングなどが含まれる．これらを検出することで，新規コマーシャルデータベースの自動構築などが可能になる．次に，(c),(d) に同じ素材映像に微小な加工が施されたものを示す．これには例えばニュース映像のような，撮影された映像に，キャプションや他の映像との合成が組み込まれたものが含まれる．短い期間に何度も繰り返し放送されるニュース映像は，重要な出来事と関連する可能性が高い．(e),(f) は同じキャストの解説映像区間である．これらを検出することで，例えば，ニュース映像の話題の区切りを見つけることが出来る．(g) はサッカーのフィールドを遠くから映したものである．

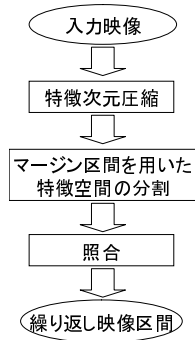


図 4 提案手法の処理手順

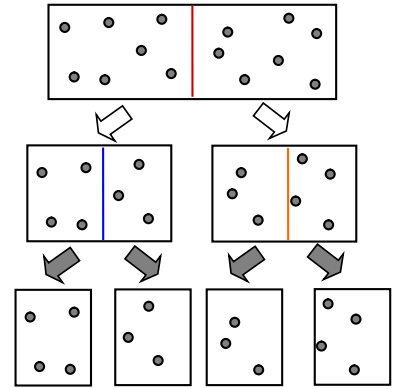


図 6 階層的分割

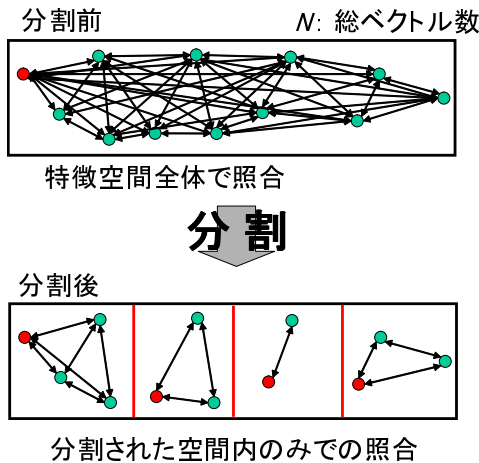


図 5 空間分割による照合回数の削減

2.2 特徴次元圧縮による高速化

従来提案されている繰り返し映像区間検出の高速化手法では、特徴次元圧縮を用いてベクトルの次元を下げ、一回の照合時間を短縮する。まず、訓練映像区間群を主成分分析することにより、次元圧縮に用いる基底、つまり固有ベクトルを計算する。次に固有値が大きいほうから、それに対応する固有ベクトルを d' 個取り出し、 d 次元の特徴ベクトルとの内積をそれぞれ取ることで、 d' 次元の低次元特徴ベクトルを得る。次元圧縮を行うことで、 d と d' の比に応じて照合時間が短縮され、各映像区間対の照合における計算の高速化が期待できる。この手法に関する詳細は文献 [4] に譲る。

3. 特徴空間のマージン付き分割による高速化

まず、文献 [4] と同様に次元圧縮によって、各映像区間を低次元特徴空間内のベクトルとして表現する。次に、この特徴空間を分割することで更なる照合時間の削減を図る。具体的には、分割境界を一つ設置して空間を 2 分割する操作を階層的に行う。このとき、分割境界付近で検出漏れが生じる場合がある。そこでマージン区間と呼ばれる領域を導入し、両側の空間を部分的に重複させることで検出漏れを防ぐ。処理手順を図 4 に示し、以下で詳しく説明する。

3.1 階層的空間分割

特徴空間を分割し、照合の対象となる特徴ベクトルを分割された空間内のものに限定することにより高速化を図る (図 5)。

分割は、次元圧縮の際の得られた固有ベクトルを軸として行う (図 8)。低次元特徴ベクトルの集合を $P = \{p_1, \dots, p_N\}$ とおく。まず、ベクトル p_i の要素のうち、 j 番目の要素に対応する値を $p_{i,j}$ とする。 $p_{i,j}$ を i 方向に見たときの中央値 m_j を算出し、これを分割境界とする。中央値を選択することで、分割後のベクトル数が均一となり、分割の効果が高くなる。 P 内の各 p_i を $p_{i,j} < m_j$ となる集合とそれ以外の集合に分割する。以上の処理を、分散 (固有値) の大きい軸 (固有値の大きい固有ベクトル) から順に分割軸を変えながら再帰的に繰り返す。その結果、階層的に分割された特徴空間を得る (図 6)。これにより、照合の対象となるベクトルが、分割された各空間内に存在するものだけに限定され、照合回数は $n_1^2 + \dots + n_x^2$ になる。ここで、 n_i は分割された空間 i に存在するベクトルの数、 x は分割された空間の数である。これにより照合回数が減少し、計算コストの大幅な削除が期待できる。

3.2 マージン区間

分割境界の付近に存在する繰り返し映像区間のベクトル対が、それぞれ異なる空間に属する場合には、空間を分割することによって、検出漏れが生じる。この問題を解決するため、分割時に精度保証を実現するための処理を導入する。分割境界を中心とした幅 θ の区間を設定し、 $|p_{i,j} - m_j| \leq \theta$ となる場合には、両方の空間に p_i を所属させるように、部分的に重複した 2 つの空間に分割する (図 7)。この重複した領域をマージン区間と呼ぶ。 θ は繰り返し区間検出の際の距離の閾値とする。これにより任意の p_i に対する繰り返し映像区間は、必ず同じ空間内に属することが保証される。このとき異なる空間に共通して含まれるマージン区間内の特徴ベクトルを共通ベクトルと呼ぶ。

共通ベクトルは 2 つの空間に所属しているため、重複した照合が起これ、分割の仕方によっては照合回数が増加してしまう場合が考えられる。このことについては後の実験で考察する。

4. 実験

提案手法の有効性を示すため、実際の映像に対して繰り返し映像区間検出を行い、空間分割後の照合時間と分割前の照合時間を比較した。

4.1 実験条件

実験では、入力映像としてニュース映像とサッカー映像を用

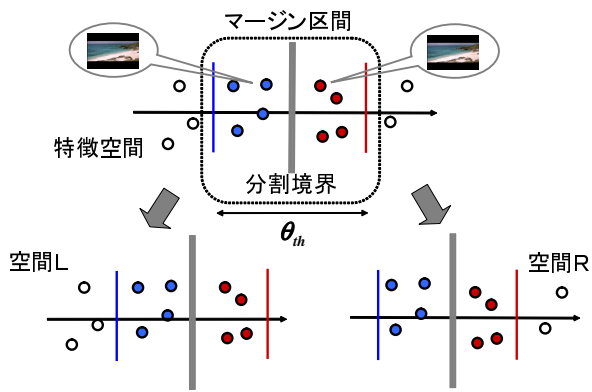


図7 マージン区間による精度保証

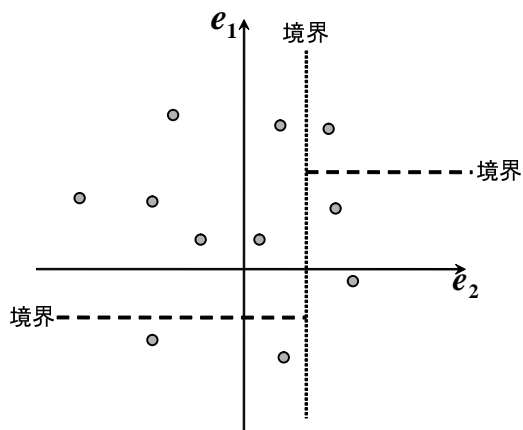


図8 分割境界の設置

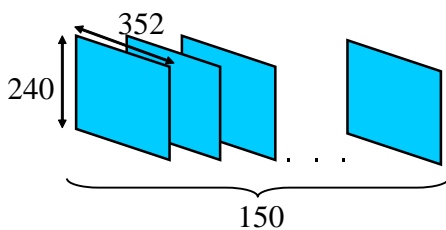


図9 検出対象となる映像区間情報

いた。ニュース映像は多様な映像表現が用いられており、特徴ベクトルの分布形状の偏りが比較的小さいのに対し、サッカー映像は芝のフィールドが大きく映りこんだ類似映像が多く出現し、特徴ベクトルの分布に大きな偏りが存在することが想定される。検出対象とする映像区間長は150フレーム(5秒)とし、特徴次元圧縮では $352 \times 240(\text{pixel}) \times 3(\text{RGB}) \times 150(\text{frames})$ の映像区間(図9)の特徴ベクトルを20次元に圧縮した。ここで、空間分割を行った際の照合時間は、分割された各空間を照合した時間の合計とした。

4.2 結果と考察

4.2.1 検出された映像区間

検出された映像区間は、同一素材のニュース映像、コマーシャルに加えて、同じキャストの解説映像区間、空や海、サッカーフィールドなどの淡色に染まった部分が大きく写された区間などであった。

4.2.2 照合時間比較

図10, 11, 12はそれぞれ60分, 180分, 300分の入力映像において、空間分割を行った際の照合時間、共通ベクトル数を分割数毎に示している。分割数1の場合が空間分割前の照合時間に相当する。いずれの実験においても、分割前と比較して、検出の高速化が確認できる。例えば、60分のニュース映像の場合、分割しない場合には638秒であったのに対し、適切に分割することにより331秒になった。

ニュース映像とサッカー映像を比較すると、図11, 12にてサッカー映像の共通ベクトル数が大幅に増えていることがわかる。これは、サッカー映像の特徴ベクトルの分布に偏りがあったためであると思われる。

またいずれの場合も、階層的に空間分割を繰り返していくと、照合時間が減少していくが、ある分割数を超えると逆に照合時間が増加する現象がみられた。これは、マージン区間の導入によって生じた共通ベクトルの数が過剰に増え、重複した無駄な照合が多く行われたためであると思われる。

4.2.3 分割数に関する検討

表1に各入力映像において、照合時間が最小となる分割数を示す。表の分割数は特徴空間を同じ深さまで階層的に2分割しながら照合時間を測定し、最も照合時間が小さくなるときの、分割された空間の数を示している。

これらの結果から、入力映像によって、照合時間が最小となる分割数にばらつきが生じていることが分かる。入力映像長が増えると最適な分割数が増える一方で、同じ映像長でもサッカーとニュースでは異なる場合も見られた。これは、入力映像によって空間内ベクトルの分布が異なることが原因の一つとして考えられる。このため実際に繰り返し区間検出を行う際には、何らかの方法を用いて入力映像に応じた分割数を決定する必要がある。図10, 11, 12の共通ベクトルの変化から、共通ベクトルの数が急激に増えるときに照合時間が増加し始めている。そのことから、共通ベクトル数に着目した、分割の打ち切りが有効であると予想される。

5. まとめ

本研究では、マージン付き分割を導入した繰り返し映像区間検出の高速化手法を提案した。実験では、放送映像を入力とし、提案手法による空間分割後と前での照合時間を比較した。その結果、照合時間の高速化が確認され、提案手法の有効性が確認が示された。

今後の課題としては、入力映像に応じた分割の打ち切りに関する検討、並列計算機を用いた、各分割空間の同時照合処理による照合時間の測定などが挙げられる。

謝辞 日頃より熱心に御討論頂く名古屋大学村瀬研究室諸氏に感謝する。本研究の一部は、文部科学省21世紀COEプログラムおよび科学研究費補助金による。また、本稿の実験では、映像情報処理にMISTライブラリ(<http://mist.suenaga.m.is.nagoya-u.ac.jp>)を使用した。

文 献

- [1] 木村昭悟, 柏野邦夫, 黒住隆行, 村瀬洋: “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, Dec. 2002.
- [2] D. Dementhon and D. Doermann: “Video retrieval using spatio-temporal descriptors”, Proceedings of ACM Multimedia 2003, pp.508-517, Nov. 2002.
- [3] X. Naturel and P. Gros: “A fast shot matching strategy for detecting duplicate sequences in a television stream”, Proceedings of The Second International Workshop on Computer Vision Meets Databases, pp.21-27, Oct. 2005.
- [4] 野田和広, 高橋友和, 井手一郎, 目加田慶人, 村瀬洋: “適応的特徴選択を用いた長時間放送映像からの高速な繰り返し区間検出”, 信学技報, PRMU2005-289, Mar. 2006.
- [5] 小川晃, 野田和広, 高橋友和, 井手一郎, 村瀬洋: “画像情報を用いた同一ニュースイベントの言語横断検索”, 第3回デジタルコンテンツシンポジウム, 2-1, June 2007.
- [6] 社本祐司, 高橋友和, 井手一郎, 村瀬洋: “同一映像区間を手がかりとしたニュース映像アーカイブのトピック構造解析”, 第3回デジタルコンテンツシンポジウム, 3-4, June 2007.
- [7] S. Arya, D. M. Mount, N. S. Netanyahu, R. Silverman, and A. Y. Wu: “An optimal algorithm for approximate nearest neighbor searching”, Journal of the ACM, Vol.45, no.6, pp.891-923, Nov. 1998.
- [8] 荒井英剛, 武本浩二, 加藤丈和, 和田俊和: “階層的固有空間による高次元最近傍探索の高速化”, 画像の認識・理解シンポジウム (MIRU)2006 論文集, pp.291-297, July 2007.

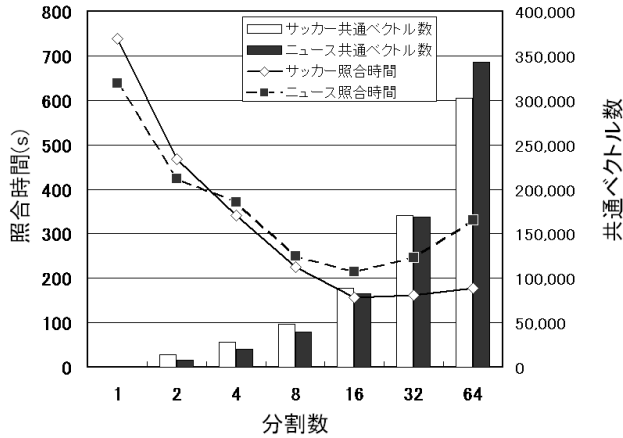


図 10 映像 60 分の繰り返し映像区間の照合時間

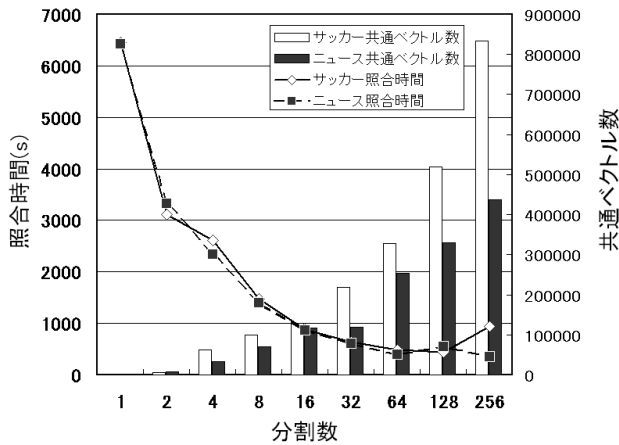


図 11 映像 180 分の繰り返し映像区間の照合時間

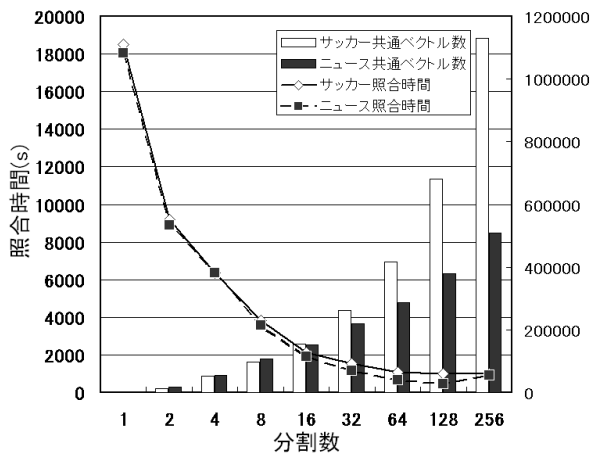


図 12 映像 300 分の繰り返し映像区間の照合時間

表 1 各映像における照合時間が最小となる分割数

ジャンル	映像時間	分割数
サッカー	60 分	16
サッカー	180 分	128
サッカー	300 分	128
ニュース	60 分	16
ニュース	180 分	64
ニュース	300 分	128