放送映像における準同一映像区間の出現パターンによる分類

社本裕司 †， 出口大輔 †， 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　†† 岐阜聖徳学園大学 経済情報学部　‡ 国立情報学研究所

近年，準同一映像区間を利用した放送映像の利用支援技術が注目を集めている．これらを利用するとコマーシャル映像データベースの構築，関連するニュースストーリの対応付け，番組の構造解析といった映像の利用支援が可能になる． 本研究では，放送映像中に存在する準同一映像区間をコマーシャルや番組のコーナを示す映像のような種類に応じて分類する手法を提案する．提案手法では、準同一映像区間の出現パターンが種類に応じて異なるという特性に注目して分類する．6 放送局で放送された 1 週間分の映像を用いた実験により，出現パターンによって準同一映像区間を分類することが可能であり，従来の種類ごとに特化した準同一映像区間の応用技術を統一的に扱える可能性を確認した．

Classification of Near-Duplicate Video Segments in Broadcast Video by Occurence Patterns

Yuji SHAMOTO †　Daisuke DEGUCHI †　Tomokazu TAKAHASHI ††　Ichiro IDE †,‡　Hiroshi MURASE †

† Graduate School of Information Science，Nagoya University

†† Faculty of Economics and Information, Gifu Shotoku Gakuen University

‡ National Institute of Informatics

Recently, technology to support the access to broadcast video using near-duplicate video segments has gained attention. By using near-duplicate video segments, it become possible to support the access to broadcast video contents such as construction of a commercial video database, tracking related news stories, structure analysis of a TV program. This paper proposes a classification method of near-duplicate video segments in broadcast video according to their usage. The proposed method refers to occurrence patterns to analyze and classify the categories of near-duplicate video segments. By an experiment with broadcast video from 6 stations for a week, near-duplicate video segments were automatically classified by the proposed method. The method presented the capability to generally handle individual application technology using near-duplicate video segments.

1　はじめに

近年，情報通信技術の進歩や記憶装置の性能向上に伴い，家庭用のハードディスクレコーダでも数百時間の映像を記録することが可能になった．また，YouTube[1]やニコニコ動画[2]等の動画共有サイトの出現により，我々が日常的に利用できる映像の量も爆発的に増大している．なかでも放送映像は非常に身近なものであり，構造解析 · 要約 · 索引付け等により映像を効率的に利用する技術や利用者が見たい映像を検索する技術が注目を集めている．

これらの技術に関して，従来のテキスト情報を中心とする手法に対して，画像情報を利用した研究が行われはじめている．その中でも準同一映像区間 (near-duplicate video segments) を利用した研究が注目されている．準同一映像区間とは放送映像の中に複数回出現する同一あるいは非常に類似した映像区間のことである．その代表的なものとしてはコマーシャル映像，同一素材を用いたニュース映像，番組コーナの先頭に使われる定型的な CG 映像などが挙げられる．これらに関して，コマーシャルデータベースの自動構築 [1)]，関連するニュースの対応付け [2) 3)]，ニュース映像の自動要約 [4)]，番組の構造解析 [5) 6)] など，さまざまな応用が期待されている．

しかし，準同一映像区間を利用する際に問題となるのは，コマーシャル映像はコマーシャルデータベースの構築に，またニュース映像は関連するニュースの対応付け

[1] http://www.youtube.com/
[2] http://www.nicovideo.jp/

に応用するといったように，種類によって応用方法が異なる点である．従来の研究では，ある特定ジャンルの映像から準同一映像区間を検出したり，特定の種類の準同一映像区間のみを検出し，暗黙のうちにその役割を仮定した上で応用方法が検討されてきた．しかし対象となる映像のジャンルや準同一映像区間の種類毎に特化した手法を考えなければならないという問題点があった．

本研究では，映像のジャンルを限定せずにあらゆる準同一映像区間を検出して，種類ごとに分類することを目指す．映像のジャンルを限定しない利点は 2 つある．1 つ目は準同一映像区間を利用する際に，これまではジャンルごとに個別に扱っていたものを統一的に扱うことができるという点である．2 つ目は，あるジャンルの映像だけでは準同一映像区間の種類がわからない場合に，他のジャンルの映像における使われ方を合わせて見ることで分類が可能になることがある点である．例えば，ニュース番組中に出現する準同一映像区間は，ニュース番組だけをみると同一素材のニュース映像である可能性が高いが断定することはできない．しかし，あらゆるジャンルの映像を見ることで，ジャンルを問わず多数出現していればコマーシャル映像であるといったように正確な分類が可能となる，このように，種類ごとに出現パターンが異なる特性に注目して，準同一映像区間の分類を行う．

以降，2 節で準同一映像区間の検出手法について述べる．3 節では準同一映像区間を種類ごとに分類する手法を説明する．4 節では実際の放送映像を用いた評価実験と考察について記す．最後に 5 節でまとめる．

2　準同一映像区間の検出

本研究では Noda らの手法 [7] を改良した手法を用いて準同一映像区間の検出を行った．この手法では，照合に用いる特徴次元を時間方向と空間方向の 2 段階に圧縮し，低次元の特徴で検出候補を絞り込むことによって，計算量を削減する．この手法に対して独自に改良した点を以下で述べる．

2.1　照合領域の切り出し

フレーム画像から照合に用いる領域の切り出しを行う．一般の放送映像では放送局が素材映像に独自のテロップを挿入するなどの編集が頻繁に行われる．画像中にテロップが存在すると，同じ素材映像を用いている映像であっても，画像的な類似度が低下する．そのためテロップが挿入されにくいと思われる中央の領域を切り出し，照合に利用する (図 1)．この領域は経験的に決めたものである．

2.2　放送局毎の色調の補正

完全に同一の映像であっても放送局により色調に違いが生じる問題がある．図 2 に示す例から，放送局 A で放送された映像は放送局 B で放送された映像と比べて輝度が低いことがわかる．このような色調の違いは放送局によって一定の傾向が見られる．そこでこれを補正するために，同一のコマーシャル映像を利用して補正式を算出する．

図 1: 照合領域の切り出し (単位: 画素)

図 2: 放送局による色調の違いと色調補正の結果

図 3: 放送局間の色相関と補正直線の例

本研究では，関東地方で放送された放送局 6 送局

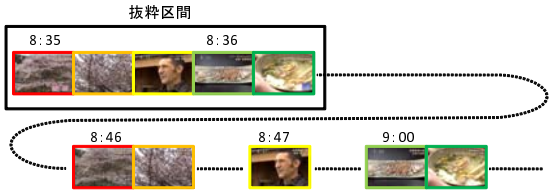

図 4: 「**抜粋区間**」の例

2.3 ラベリング

照合によって多数の準同一映像区間対が得られるが，同一の映像が関連付く対がある．そこで，準同一映像区間のラベリングを行う．以降，同じラベルが与えられた準同一映像区間の集合を準同一映像区間群を呼ぶ．

3 準同一映像区間の分類

3.1 準同一映像区間の分類種別

本研究では，応用を考慮して準同一映像区間を以下の 6 種類に分類する．

1) **コマーシャル**: コマーシャル映像
2) **ニュース**: 同一素材映像を用いたニュース映像
3) **コーナ**: オープニングやコーナの先頭 CG 映像等，番組の構成を表すのに使われる同一映像
4) **類似構図**: キャスタショット等，構図が類似した映像
5) **再放送**: 番組の再放送
6) **抜粋区間**: 別の映像から抜粋した映像で構成される区間

6) は単体でみると単なる準同一映像区間であるが，図 4 のように別の時刻に出現する映像から短い区間を抜粋して連結した映像のことである．具体例としてある番組のオープニングで番組内容を紹介する映像のほか，番組の次回予告，スポーツのハイライト映像がある．

上記の各種類の準同一映像区間について，以下の応用方法が考えられる．

1) **コマーシャル**: コマーシャル映像の除去やコマーシャル映像データベースの構築
2) **ニュース**: 関連するニュースストーリの対応付け
3) **コーナ**: 番組の構造解析や特定コーナの抽出
4) **類似構図**: 番組の構造解析に使えるが，応用によっては不必要
5) **再放送**: シリーズものの番組でどの回の再放送であるかの判別
6) **抜粋区間**: 本編と対応付けて利用者に提示

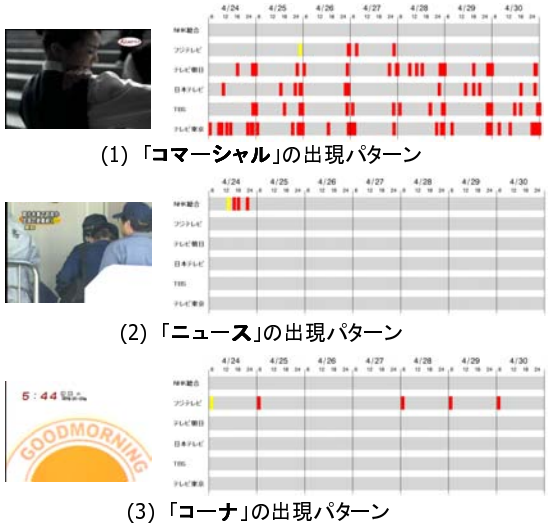

図 5: 準同一映像区間の出現パターン

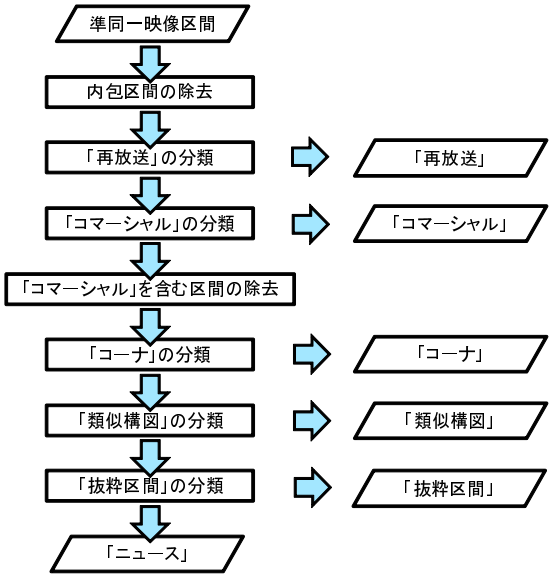

図 6: 準同一映像区間の分類手順

3.2 分類手順

図 5 に示すように，種類によって出現パターンの傾向が大きく異なる．そこで提案手法では準同一映像区間の出現パターンに関する条件を人手で設定し，準同一映像区間の時間長と合わせて分類を行う．分類手順を図 6 に示す．図に示すようにクラスごとの条件を満たす準同一映像区間を順に抽出していくことによって分類を行う．

分類の前に，ある準同一映像区間群が別の準同一映像区間群を完全に内包する場合 (図 7)，内包された区間は除去する．

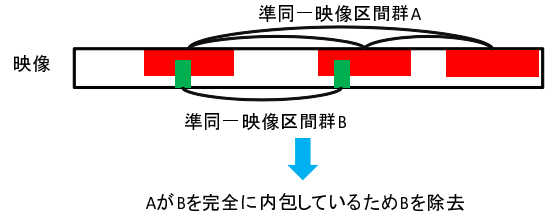

図 7: 内包区間の除去

まず,「**再放送**」と「**コマーシャル**」に関しては準同一映像区間自身の時間長が特徴的である．よって，以下のように分類する．

「再放送」の分類

「**再放送**」はある 1 つの番組全体あるいはコマーシャルで区切られた長い映像が準同一映像区間として検出される．そこである準同一映像区間群のうち θ_{time_error} の割合で区間の時間長 t [フレーム] が θ_{repeat_time} 以上のとき「**再放送**」に分類する．

「コマーシャル」の分類

「**コマーシャル**」は一般に時間長が決まっている．日本では 15 秒 (450 フレーム)，30 秒 (900 フレーム)，60 秒 (1800 フレーム) のいずれかであるものがほとんどである．そこで「**コマーシャル**」を分類する際に準同一映像区間の時間長を用いる．若干の変動やコマーシャル映像の前後に類似した映像がある場合を考慮して，ある準同一映像区間群のうち θ_{time_error} の割合で区間の時間長 t[フレーム] が $450-t_{cm_time_error} < t < 450+t_{cm_time_error}$，$900-t_{cm_time_error} < t < 900+t_{cm_time_error}$，$1800-t_{cm_time_error} < t < 1800+t_{cm_time_error}$ のいずれかの条件を満たすとき「**コマーシャル**」の候補とする．しかし，この条件だけでは偶然 15 秒，30 秒，60 秒の時間であった「**コマーシャル**」以外のものを除去できない．そこで，ある準同一映像区間群内の区間が出現している期間が θ_{cm_span} 日以上である，あるいは出現している放送局数が 2 局以上であるという条件を与える．

ここで，あるコマーシャル映像と別のコマーシャル映像の一部が共通する場合に共通部分が準同一映像区間として検出される．また，同じ順番でコマーシャル映像が連続して放送される場合に複数のコマーシャル映像が 1 つの準同一映像区間として検出される．そこで,「**コマーシャル**」として分類された準同一映像区間群と一部でも重複を持つような準同一映像区間群をコマーシャルの一部やコマーシャルの連結として分類から除外する．

「コーナ」の分類

次に「**コーナ**」の分類を行う．図 8 に処理の概要を示す．分類実験では 1 週間分の放送映像を対象にするので，毎日放送される番組で出現する準同一映像区間を考える．そこで周期的に出現するという条件を毎日同じ時間帯に出現するという条件に置き換える．厳密に同じ時間に出現するとは限らないため時間帯の幅として L_{corner} の間に収まるとき同じ時間帯に出現するものとみなす．また，出現する時間帯が複数ある場合も許容する．また，番組が放送されない日や特定のコーナがない日を想定して，毎日ではなく θ_{corner_date} 日以上出現していることを条件として与える．さらに，若干の誤検出を許容するために，区間のうち θ_{corner_error} の割合で時間帯と出現日数の条件を満たす場合に「**コーナ**」と分類する．

図 8: 「**コーナ**」の分類

「類似構図」の分類

続けて「**類似構図**」の分類を行う．まず，出現する放送局数が 1 局であるという条件を与える．また，非常に短い間隔で出現する場合に「**類似構図**」であると考える．ただし，複数日の同じ番組で類似した構図がある場合もあり，日をまたぐ場合には間隔があく．そこで，出現間隔の中央値が $\theta_{interval}$ 以下という条件を与えることで間隔があく場合に対処する．

「抜粋区間」の分類

「**抜粋区間**」は別の映像から抜粋してきた映像区間を構成する複数の準同一映像区間群をまとめて分類する.「**抜粋区間**」の分類について図 9 に示す．まず，放送局数が 1 局であるという条件を与える．その中で準同一映像区間が連続している部分を抽出するために，複数の準同一映像区間の時間差が $\theta_{abstract_space}$ 以下で，かつ $\theta_{abstract_number}$ 個以上接して出現する部分を抽出する．抽出された準同一映像区間の集合について接する部分以外を見た時にある時間幅 $\theta_{max_abs_range}$ の中で全種類の準同一映像区間が出現して，かつある時間幅 $\theta_{min_abs_range}$ では全種類が出現しない場合に「**抜粋区間**」と分類する．

「ニュース」の分類

「**ニュース**」に関してはさまざまな出現パターンが考えられ，分類の条件がを決められなかったため，最後に残された準同一映像区間をすべて「**ニュース**」と分類

図 10: 「**抜粋区間**」の分類誤り例

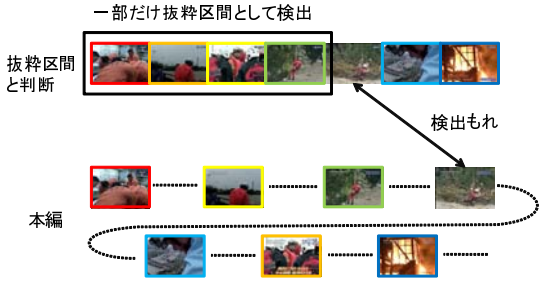

図 11: 「**抜粋区間**」の一部のみが検出される例

番組においてコマーシャルでない映像の前後で出現するリプレイ映像が多く存在した．現状では類似構図を分類する際に出現の間隔しか見ていないが，その他の出現パターンを用いることで精度が向上できると考えられる．

「**抜粋区間**」には，図 10 のように長いニュース番組内などで同じトピックの短いニュースを何回も放送する場合に，内容の微妙な違い等によって抜粋区間と判定される場合が多かったため精度が悪かった．一般にこのようなニュース映像は数多く出現すると考えられるので，出現頻度の上限に関する条件を与えるなどの解決策が考えられる．また，抜粋したと思われる区間の一部しか「**抜粋区間**」と検出されない場合が多かった．これは図 11 のように本編に出現するが準同一映像区間検出からもれた映像や，そもそも本編には出現しない映像が間に挟まったために生じたと考えられる．「**抜粋区間**」を抽出する際に，検出もれや余分な映像の存在に頑健な手法を検討する必要がある．

「**ニュース**」に関しては，現状では明確な出現パターンの傾向を見つけることができなかったため，どのクラスにもあてはまらない準同一映像区間が数多く分類された．また，「**抜粋区間**」の分類手法が十分でなかったために，人手による分類では「**抜粋区間**」の区間群を誤分類する結果となった．改善策としては，「**ニュース**」を特徴づける出現パターンの検討や新たなクラスを追加する手法が考えられる．

5　まとめ

本報告では，あらゆるジャンルの放送映像から準同一映像区間を検出し，コマーシャル映像やコーナの先頭を表す定型的な映像などの種類ごとに自動分類する手法を提案した．実験の結果，「**コマーシャル**」や「**コーナ**」は高い精度で分類できることを確認した，また，その他のクラスについてもある程度の精度で分類できたので，準同一映像区間の自動分類により，映像のジャンルや準同一映像区間の種類を限定した従来の利用支援技術を統一的な枠組みで行える可能性を示した．

今後の課題としては，自動分類手法の改良として機械学習による分類規則の学習や電子番組表 (EPG) の情報を併用した分類の検討，より大規模な映像に対する適用が挙げられる．

謝辞

日頃から熱心に御討論頂く村瀬研究室の諸氏に深く感謝する．本研究の一部は科学研究費補助金特定領域研究「情報爆発 IT 基盤」及び国立情報学研究所との共同研究による．本研究では，画像処理に MIST ライブラリ (http://mist.suenaga.m.is.nagoya-u.ac.jp/) を使用した．

参考文献

1) Y. Li, D. Zhang, X. Zhou, and J. S. Jin: “A confidence based recognition system for TV commercial extraction”, Proc. 19th Australasian Database Conf., pp.57–64 (2008)

2) 社本裕司，高橋友和，井手一郎，村瀬　洋: “大規模ニュース映像アーカイブにおける繰り返し映像区間の出現傾向分析”，情報処理学会第 70 回全国大会講演論文集，Vol.5，pp.217–218 (2008)

3) 井手一郎，小川　晃，高橋友和，村瀬　洋: “画像とテキストの利用による同一ニュースの言語横断検出”，画像ラボ，Vol.19，No.12，pp.22–26(2008)

4) X. Wu, C-W. Ngo, and Q. Li: “Threading and autodocumenting news videos”, IEEE Signal Processing Magazine, Vol.23, No.2, pp.59–68 (2006)

5) X. Yang, Q. Tian, and P. Xue: “Efficient short video repeat identification with application to news video structure analysis”, IEEE Trans. Multimedia, Vol.9, No.3, pp.600–609 (2007)

6) 青木　恒: “映像対話検出によるテレビ番組コーナ構成高速解析システム”，電子情報通信学会論文誌, Vol.J88-D-II，No.1，pp.17–27 (2005)

7) I. Ide, K. Noda, T. Takahashi, and H. Murase: “Genre-adaptive near-duplicate video segment detection”, Proc. 2007 IEEE International Conference on Multimedia and Expo, pp.484–487 (2007)