

大量ニュース映像を対象とした時系列意味構造に基づく情報編纂手法の提案

Proposal of an Information Compilation Method for Massive News Video Data Based on their Time-Series Semantic Structure

井手 一郎

Ichiro Ide

名古屋大学 大学院情報科学研究科 / 情報・システム研究機構 国立情報学研究所

Graduate School of Information Science, Nagoya University / National Institute of Informatics, ROIS
ide@is.nagoya-u.ac.jp / ide@nii.ac.jp

木下 智義

Tomoyoshi Kinoshita

(株) ネットコンパス

NetCOMPASS Ltd.
kino@netcompass.co.jp

高橋 友和

Tomokazu Takahashi

岐阜聖徳学園大学 経済情報学部

Faculty of Economics and Information, Gifu Shotoku Gakuen University
ttakahashi@gifu.shotoku.ac.jp

孟 洋

Hiroshi Mo

情報・システム研究機構 国立情報学研究所

National Institute of Informatics, ROIS
mo@nii.ac.jp

片山 紀生

Norio Katayama

(同 上)

katayama@nii.ac.jp

佐藤 真一

Shin'ichi Satoh

(同 上)

satoh@nii.ac.jp

村瀬 洋

Hiroshi Murase

名古屋大学 大学院情報科学研究科

Graduate School of Information Science, Nagoya University
murase@is.nagoya-u.ac.jp

keywords: video archive, topic tracking, topic thread structure, browsing interface

Summary

Recent increase of digital storage capacity has enabled the creation of large-scale on-line broadcast video archives. In order to make full use of the data in the archive, it is necessary to let a user easily grasp the availability of certain video data and their contents. Considering this problem, we have been investigating efficient and effective retrieval and reusing methodologies of archived video data. The archive used as a test-bed consists of more than 1,000 hours of news video obtained from a Japanese news program during the past six years. This paper first proposes a news topic tracking and structuring method. A structure called the 'topic thread structure', is organized so that it should represent the temporal flow of news topics originating from a specified news story. The paper next introduces a browsing and editing interface that enables the user to browse through news stories along the topic thread structure, and also assists the compilation of selected news stories as a customized video summary or a documentary. The method was applied to the archived news video data in order to observe the quality of the topic thread structure and the usability of the prototype interface. As a result, some structures represented the flow of topics quite close to real-world comprehension. In addition, experiments showed that when the structure could be considered meaningful, the interface combined with the structure could drastically reduce the time needed to browse through the archive for news stories related to the user's interest.

1. はじめに

1.1 背景と目的

近年の記憶装置の大容量化に伴い、大量の放送映像をオンラインで蓄積して利用できるようになりつつある。なかでもニュース映像は、実世界に生起する様々な事象の時系列に沿った記録であるため、蓄積したものを検索・再

利用する価値が高い。このような観点から、我々は2001年より6年以上にわたり特定のニュース番組の映像を連日収録し [Katayama 04], 1,000 時間を超える映像の効率的かつ効果的な検索・再利用方法を検討してきた。

図1に示すように、日々放送されるニュース映像中の1つ1つの話題(ストーリー; 同図左側)は、一般に数日にわたって継続するトピックの流れ(トピックスレッド; 同

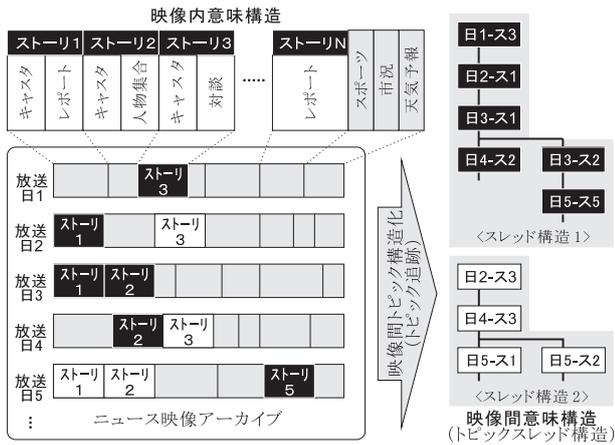


図 1 ニュース映像の意味構造 .

図右側)の断片である．そのため，各ストーリーの個別内容だけでなく，トピックとして内容を総合的に理解することが重要である．しかし，任意の事象に関するニュースについて，意味内容に則して関連付け，追跡・理解するためには，日々蓄積される膨大なニュースの内容の全貌を把握する必要がある．このような作業は人間の能力を越えるため，計算機による支援が不可欠である．

そこで我々は本論文において，まず大量のニュース映像から，意味的に強く関連するストーリーを時系列に連鎖した構造：トピックスレッド構造を抽出する手法を提案する．次に，ニュース映像中の指定した1つのストーリーを起点とするトピックの変遷をたどり，その顛末をまとめた要約映像の編纂を支援するために試作したインタフェース：mediaWalkerを紹介する．

1.2 概念と用語の定義

ニュース映像や新聞などニュース一般に固有の概念について定義する．本論文では，新聞記事を対象とした話題の追跡を目指して米国立標準技術研究所 (NIST) が主催したワークショップ TDT (Topic Detection and Tracking) における定義 [Nat 00] を準用する．

- イベント：いわゆる 5W1H のうちの 4W (When, Where, Who, What) で表現される，特定の事象．
- ストーリー：一つのイベントを扱った，新聞記事やニュース映像の意味的な最小単位．
- トピック：共通のイベントに関連して強く関連するストーリーの集合．

一方，次の用語は，本論文において独自に定義する．

- トピックスレッド：強く関連するストーリーを時系列に次々に連鎖したもの．経路上に複数のトピックを含むこともある．
- トピックスレッド構造：同一のストーリーを起点とするトピックスレッドを束ねた有向グラフ構造．
- トピッククラスタ：トピックスレッド構造中の隣接するストーリーの意味的まとまり．

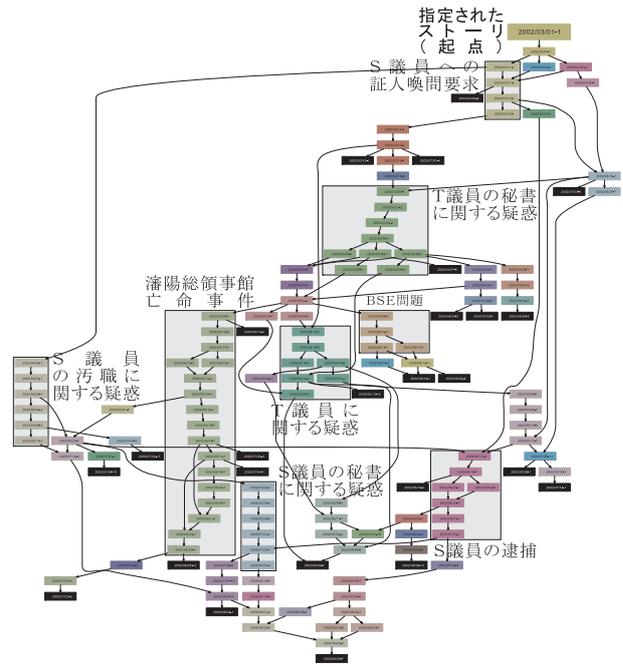


図 2 大規模なトピックスレッド構造の例．2002年3月1日冒頭から1番目のストーリーを起点とする構造．探索期間 $d = 180$ [日]，類似度閾値 $\theta_{trk} = 0.40$ ．計算時間：845 [秒]．

1.3 提案手法の概要

まず，本研究で提案するトピックスレッド構造について，簡単にその構築方法と，得られる構造の例を紹介する．

前節で定義したようなトピックスレッド構造を得るために，時系列的前後関係を維持しながら関連するニュースストーリーを接続する．その結果，各エッジにおいて局所的な関連性（意味的な関連性と時系列上の前後関係）が表現されるだけでなく，起点のストーリーから派生するトピックの分岐・合流といった大きな流れも表現される．

図 2 に提案手法により得られた大規模なトピックスレッド構造の例を示す．4 ストーリー以上からなるトピッククラスタを重ねて表示し，その内容を人手で確認したものを併記した．この例から，「S 議員への証人喚問要求」が「S 議員の逮捕」という直接関連しそうなトピックだけでなく，「T 議員に関する疑惑」や「瀋陽総領事館亡命事件」など直接の因果関係が自明でないトピックへの変遷を見てとることができる．このような構造をたどることで，いわば「風が吹けば箱屋がもうかる」というような，一見関連が分からない実世界の事象間について理解することを支援したい．

このような構造は，映像閲覧インタフェースとともに使われてこそ効果を発揮する．そこで，トピックスレッド構造に基づき映像を閲覧するインタフェースを試作した．このインタフェースはさらに，トピックスレッド構造上の指定した映像を選択してプレイリストとして出力する機能を備えているため，利用者が映像アーカイブの全体を把握していなくても，ある事象間の関連を説明するために，トピックスレッド上の映像を順に接続して編

纂することができる。

以降、2 章において関連研究を紹介した後、3 章において大量のニュース映像から時系列意味構造を分析する手法を提案し、続く 4 章において実際のニュース映像から得られた時系列意味構造の内容や規模について分析することで、提案手法の効果を検討する。さらに、5 章において、提案手法によって得られた時系列意味構造に基づいてニュース映像を閲覧しつつ、要約映像の編纂を支援する mediaWalker インタフェースを紹介する。最後に 6 章において今後の課題を検討し、本論文をまとめる。

2. 関連研究

2.1 ニュースの時系列意味構造解析に関する研究

本研究と同様の問題意識に基づき、ニュースの時系列意味構造を解析する既存手法を紹介する。

最も単純な手法として、特定のトピックに関連するストーリーを集めたクラスタにおいて、関連が強いと判断されたストーリーを時系列に直線状に連ねる方法が考えられる。これは、TDT ワークショップにおいて定義されたトピック追跡 [Nat 00] の考え方である。このような方法に基づきニュース映像を対象としたトピック追跡手法として、たとえば、Duygulu らによる同一画像・ロゴの出現を手がかりとしたもの [Duygulu 04] がある。しかし、この方法では大規模なニュースアーカイブを対象として利用者が話題を追跡・理解したい場合に、膨大なデータを直線状に逐一たどらなければならないため、実用性に問題がある。さらに直線構造では同一トピックであっても同時並行して進む個別の話題の流れを表現できない。TDT ワークショップの最終回では前者の問題の解決につながる階層的トピッククラスタ抽出という課題が導入されたが、直線状に並んだデータを逐一たどらなければならない点は同様である。

これに対し Wu らは、特定のトピックに関連するストーリーを集めたクラスタにおいて、時系列的前後関係と話題の変化（新規ストーリーの出現）に応じて 2 分グラフを構築する手法を提案した [Wu 06]。しかしこの方法でも、新規ストーリー同士の関係は時系列的前後関係のみであるため、同時並行して進む個別の話題の流れを表現できない問題が残る。

そこで我々は、同時並行して進む話題の流れを表現する時系列意味構造：トピックスレッド構造を抽出する手法を提案する。同様の取組みとして、新聞記事を対象とした研究 [Nallapati 04, Uramoto 98, Yang 06] があるが、話題の分岐点におけるトピックスレッド間の独立性が考慮されていないため、得られるグラフ構造をたどる際に冗長なエッジが出現してしまう（3.3 節で詳述）。

一方、上記の既存手法の多くは事前に同一トピックのクラスタを作成することで、追跡対象ストーリーを限定し

ている*1。これらの手法は、徐々に変遷してやがて異なるトピックへと発展するニュースの一般的性質が考慮されていない。大規模なニュースアーカイブを対象とする場合、図 2 に示した例のように、直接関係がない実世界の事象間の関係を理解することこそが重要なため、このような従来の枠組では不十分である。

これに対して我々は、事前にトピックの範囲を限定することなく、まず徐々に変遷する話題の流れをトピックスレッド構造として抽出した後に、その構造中の局所的な意味的まとまりを抽出する手法を提案する。

2.2 ニュース映像の可視化に関する研究

次に、映像の可視化に関する既存研究を紹介する。映像を静止画像により可視化する一般的手法は、映像中の任意枚数のフレームを抽出して代表画像として提示する、いわゆる storyboard と呼ばれるものである。これには先頭フレームや一定間隔でサンプリングしたフレームを表示するものと、画像中の動きや顔の出現などの特徴に基づきフレームを抽出するものがある。これらの手法は、静止画像による可視化のほか、本来長い映像を短い要約映像として提示するための研究にも用いられる。これに対して谷口らは、代表フレームを選択せずに映像を可視化する PanoramaExcerpts という手法を提案した [谷口 99]。この手法では、ショット中の各フレームからパノラマ画像を生成して代表画像として用いることにより、映像中の画像内容を余すことなく可視化することに成功した。

一方、本研究と同様に、ニュース映像に関してその特性を考慮した映像可視化手法に関する研究がある。一般にニュース情報においては、いわゆる 5W1H と呼ばれる属性をもつ情報が重要とされる。このうち、特に When, Where, Who の属性に注目した研究が多い。

Christel らは、Where の属性に注目し、ニュースの発生地点を地図上に投影して、利用者が指定した地域に関するニュースを閲覧するインタフェースを提案した [Christel 00]。彼らはこれに更に When と Who の属性を加え、video collage という可視化手法を提案し、属性の組合せによりいくつかの異なる閲覧画面を提供した [Christel 02]。また、Ide らは、Who の属性にのみ注目し、アーカイブ中の映像に出現する人物の人間関係をたどって閲覧するインタフェース trackThem を提案した [Ide 05]。これらは特定の属性に注目した関連付けに基づく閲覧手法であるため、本研究で注目するトピックの流れのように、内容の総合的な関連付けに基づく手法とは性質が異なる。

一方、本研究で提案する手法と同様に、When 以外の属性を明確に分離せず、個々のストーリーに関する全体的な特徴（テキスト情報や画像情報）の関連性に基づき、時系列に閲覧する手法がある。Rautiainen らは画像特徴に基づくニュース映像クラスタ中の類似ストーリーを閲覧

*1 ただし、[Nallapati 04] は我々と問題点を共有して、事前に追跡対象ストーリーを限定していない。

する手法として, cluster-temporal browsing を提案した [Rautiainen 04]. Snoek らは, 概念の関連性と時系列性に従って, 球面上や 2 次元平面上に映像を配置するインタフェース MediaMill を提案した [Snoek 05].

以上の可視化・閲覧手法と, 本研究で試作したインタフェースとの比較については, 5.3 節で述べる.

3. ニュース映像群の時系列意味構造解析

3.1 解析処理の流れ

トピックスレッド構造を構築するために, ニュース映像に対して以下の処理を施す:

(1) ストーリの分割

まず各々のニュース番組映像を, 本研究で扱うニュースの最小単位「ストーリー」に分割する. 図 1 中の「映像内意味構造」を解析することに相当する.

(2) トピックスレッド構造の構築

次に処理 (1) において得られたストーリーについて, ストーリ間の意味的関連性・時間的前後関係を考慮して「トピックスレッド構造」を構築する. 図 1 中の「映像間意味構造」を解析することに相当する.

(3) トピッククラスタ構造の抽出

最後に処理 (2) において得られたトピックスレッド構造中の隣接するストーリーの意味的なまとまりを「トピッククラスタ」として抽出する.

これらの処理につき, 以下 3.2 節-3.4 節で詳述する.

3.2 ストーリの分割

各ニュース番組を, 意味構造を扱うための単位であるストーリーに分割する. アンカーショットの存在やテロップの出現など映像のパターンを利用したストーリー分割手法 [Zhai 05] があるが, この方法は番組構成やデザインの影響を大きく受けるため, 本論文では映像に付随して放送される文字放送字幕 (closed-caption) テキストを用いる. なお, 放送時には音声と文字放送字幕テキストの間に時間的なずれがあるが, 以下の実験ではアーカイブ蓄積時に同期してあるものを用いた.

一般的なテキストにおけるストーリー分割手法は数多く研究されているが, 本論文ではこの処理を提案することが趣旨ではないため, 比較的簡単な手法を用い, 以下に簡単に手順と性能を紹介するにとどめる. 詳細は文献 [井手 03] を参照されたい.

(1) 文字放送字幕テキストの各文に形態素解析^{*2}を適用し, 名詞列を抽出する.

(2) 文献 [井手 02] で提案した手法に従って, 名詞列の語義属性 (一般, 人物, 場所・組織, 時相) を解析する. 各文に対し, 語義属性別に 4 つのキーワードベクトルを作成する.

(3) ある文とそれに続く文が同一ストーリーに属するか判定するために, 語義属性別に両文の前後 w 文を結合したキーワードベクトル間の距離を余弦尺度で評価する. ここで, w を 1 文から 10 文まで変え, 最後に各 w における距離の最大値を類似度とする.

(4) 各語義属性別の類似度を重み付き和の形で統合し, 閾値 θ_{seg} を下回る場合に該当箇所の文と文の間にストーリー境界が存在すると判定する. 属性別に異なる重みを与えるのは, 特にニューステキストにおいて, ストーリ分割を考える際に特定の属性が他の属性よりも重要な役割を果たすと考えたためである.

以下の実験では, 実験的に得られた重み (一般, 人物, 場所・組織, 時相) = (0.23, 0.21, 0.48, 0.08) と閾値 $\theta_{seg} = 0.17$ を用いた.

(5) 分割された各ストーリーに対して再度キーワードベクトルを作り, 手順 (3) (4) 同様に隣接するストーリー間の類似度の重み付け和が閾値 θ_{cat} を上回れば再結合する. 再結合が起きなくなるまでこの操作を再帰的に繰り返す. 閾値は実験的に $\theta_{cat} = 0.08$ とした.

人手でストーリー境界を与えた評価事例 (14 日分, 合計 130 トピック) を用いて分割結果を評価したところ, 前後 1 文の誤差を許した場合, 再現率 95.4%, 適合率 90.5% のストーリー抽出性能が得られた. ストーリ境界周辺の文は短かく, 重要な情報を含まないことが多いため, 1 文程度の誤差は許容できると考えれば, この手法により実用上ほぼ問題のない精度が得られる.

3.3 トピックスレッド構造の構築

次に, アーカイブ中の指定されたストーリーを起点とするトピックスレッド構造を構築する. トピックスレッド構造構築の目的は, 強く関連したストーリーを時系列に連鎖し, 利用者が関心をもっているトピックの顛末を順を追って理解できるような「道筋」: トピックスレッドを提示することである. 利用者が大規模な映像コーパスを探索する際には, 多数の類似したストーリーから関心があるものを選択しながら追跡することが大きな負担になるため, トピックスレッド構造から 1 本のトピックスレッドを選ぶ際の選択肢を最少限に抑えることが重要である.

単純に関連するストーリーを時系列順に階層的に展開した木 (図 3 (a)) を用いる方法も考えられるが, 各ストーリーにおいて強く関連するものの候補として多数の分岐 (選択肢) が見られる. しかし, 提案手法によるトピックスレッド構造 (図 3 (b)) では, 各ストーリーは 1 回のみ, 最も深い階層のもののみを残すように構築するため, 相対的に分岐数が抑制される.

あるストーリー (S_0) を起点とするトピックスレッド構造を構築する手順を以下に示す.

(1) ストーリ間の関連性評価

ストーリー間の関連性を評価し, 強く関連すると判断されたストーリー対を連鎖する. ストーリ間の関連性

*2 日本語形態素解析システム JUMAN を使用.

<http://nlp.kuee.kyoto-u.ac.jp/nl-resource/juman.html>

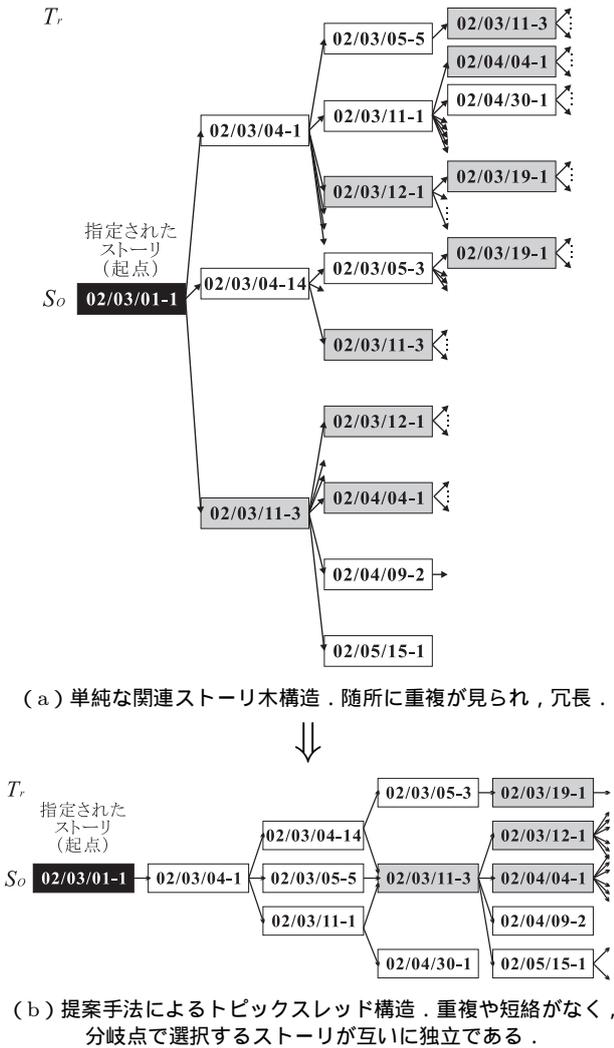


図 3 単純な関連ストーリー木構造とそれに対応するトピックスレッド構造の例．各ストーリーは [年/月/日 - 冒頭からのストーリー番号] で表記されている．灰色のストーリーは (a) において重複しているものを示す．

は，ストーリー毎に作ったキーワードベクトルの余弦距離を用いて評価し，ある閾値 θ_{trk} を上回れば強く関連しているとみなす．

(2) 単純な関連ストーリー木の構築

次の条件を満たしながら， S_0 を根とする単純なストーリー関連木 T_r (図 3 (a)) を展開：

- a 子供は親と強く関連し，必ず親よりも新しい
- b 兄弟は必ず年少の方が新しい

(3) 重複ストーリーの除去

次に， T_r 中の全ての部分木 $T_s(i)$ に対して，年長の節点に等価な部分木 $T_s(j)$ が存在するとき，次の操作を施す：

- a 除去： $T_s(j)$ が $T_s(i)$ の兄の子孫であるとき， $T_s(i)$ を除去
 - b 統合： $T_s(j)$ が $T_s(i)$ の先祖（親を除く）の子孫であるとき， $T_s(i)$ を $T_s(j)$ と統合
- 以上の除去・統合の仕組みを図 4 に示す．

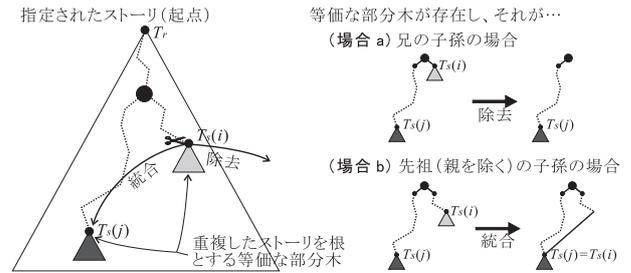


図 4 トピックスレッド構築アルゴリズムにおける操作．

ここで手順 (3) a において統合する代わりに除去することにより，分岐点の次のストーリー間に強い関連性がないことが保証される．このため，分岐点の次のストーリーを選択することが，少なくともその段階では独立したトピック(スレッド)を選択することになる．既存研究 [Nallapati 04, Uramoto 98, Yang 06] のように，統合操作のみを適用した場合，この独立性が保証されず，選択枝が冗長になり得るため，利用者に余計な選択に伴う負荷を与えるおそれがある．

ここでは未来方向のスレッド構造を構築する手順を示したが，過去方向の場合は「新しい」を「古い」に置き換えればよい．なお，起点となるストーリー S_0 は，5 章で紹介するようなインタフェース (図 7 (a)) を通じて表示される，何らかの手段により検索された結果や，指定された日に放送されたストーリーの一覧から指定する．

3.4 トピッククラスタの抽出

得られたスレッド構造をより効果的に利用するために，トピックスレッド構造中の隣接するストーリーについて，内容が大きく変わらないもののみをまとまり (トピッククラスタ) を次のようにして抽出する．

- (1) ストーリー S_0 をクラスタ中心 ($C_0 = S_0$)，かつ注目ストーリー ($S = S_0$) に指定．
- (2) 注目ストーリー S の子を $S_c(j) (j = 1, \dots, J)$ とする．このときストーリー C_0 と $S_c(j)$ との類似度 (3.3 節と同様に定義) が閾値 θ_{cls} 以上ならば， $S_c(j)$ を新たなクラスタ中心に指定 ($C_0 = S_c(j)$) ．
- (3) 全ての葉ノード $S_D(i) (i = 1, \dots, L)$ に到達するまでスレッド構造を走査して注目ストーリー S をずらしていき，手順 (2) を再帰的に適用．
- (4) 最後に，各クラスタ中心からトピックスレッド構造をたどって次のクラスタ中心までに含まれる全てのストーリーを同一トピッククラスタに属すると判断．

このような構造を得ることにより，トピックスレッド構造のなかでトピックがどのように変遷しているか，いわば話題の節目を把握しやすくなり，意味構造に基づいた情報提供を効果的に行なえと考える．

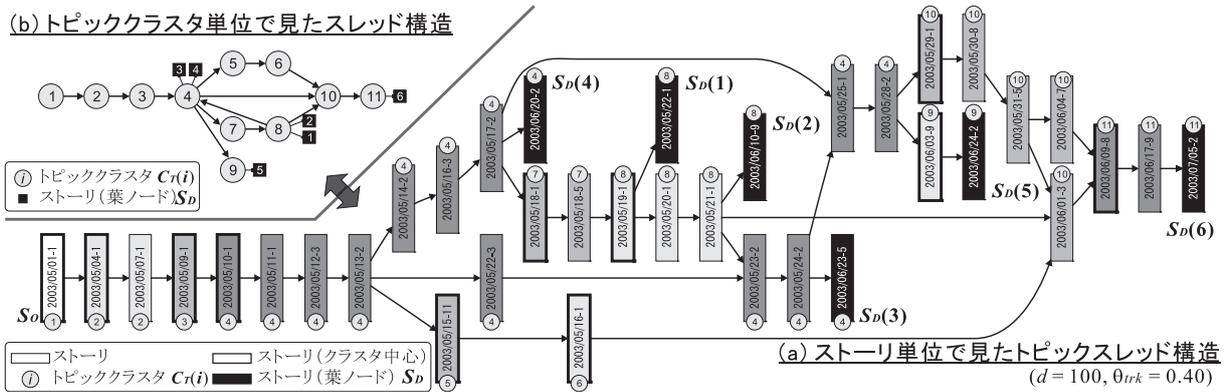


図 5 2003 年 5 月 1 日番組冒頭から 1 番目のストーリーを起点とするトピックスレッド構造 .

表 1 図 5 におけるトピッククラスタの具体的内容 .

トピッククラスタ	具体的内容
$C_T(1)$	北京で SARS 発生
$C_T(2)$	中国本土で感染拡大
$C_T(3)$	WHO が北京へ使節派遣
$C_T(4)$	中国本土でやや鎮静化, 台湾で拡大
$C_T(5)$	中国政府が地方での拡大を懸念
$C_T(6)$	中国政府が地方での拡大を監視
$C_T(7)$	日本旅行から帰国後に台湾人医師の感染発覚
$C_T(8)$	日本国内の感染調査
$C_T(9)$	北京で SARS 対策会議開催
$C_T(10)$	中国本土で鎮静化, トロントで発生
$C_T(11)$	台湾で鎮静化, WHO が終結宣言

表 2 分析に用いたトピックスレッド構造の構築条件 .

番組	NHK ニュース 7 (毎夜 20-30 分)
起点ストーリー	2002 年 3 月 1 日- 6 月 30 日
S_0 の放送日	(合計日数: 122 日)
探索期間 d^{*3}	100 日間 (映像長: 171,600 秒)
類似度閾値 θ_{trk}	0.40

ここで、構造の妥当性を定量的に評価するのは難しいため、4.2 節において解釈が困難な構造の例と原因、その対策について記すにとどめる。また、意味がある構造が得られたときに、提案手法で関連があると判断された映像のみを閲覧することによる効率化について 4.3 節において調べる。

3.5 トピックスレッド構造の例

図 5(a) に、本手法により得られたトピックスレッド構造の例を示す。また同図 (b) には同じ構造をトピッククラスタ単位で表現したものを、表 1 に各トピッククラスタの具体的内容を人手で分析した結果を示す。トピッククラスタ単位の構造により、長期にわたるトピックスレッド構造を巨視的に把握できることが分かる。たとえば、 $C_T(1)$ $C_T(2)$ $C_T(3)$ $C_T(4)$ $C_T(10)$ $C_T(11)$ とたどるトピックスレッドを選択すれば、2003 年前半に起きた SARS 騒動の顛末の本流が把握できる。また、 $C_T(4)$ $C_T(5)$ $C_T(6)$ $C_T(10)$ とたどるトピックスレッドでは中国国内における顛末が、 $C_T(7)$ $C_T(8)$ とたどるトピックスレッドの一部では日本国内へ飛び火する様子といった副次的な流れも把握できる。

4. 実験：トピックスレッド構造の分析

実際に放送されたニュース映像に以上の処理を適用し、得られたトピックスレッド構造を分析する。ここでは、表 2 に示す条件のもとでトピックスレッド構造を構築した。

4.1 実験結果

まず、ストーリー分割により、2 文以上からなるストーリーを抽出した。ここで、1 文からなるストーリーは、過分割による断片的な短文であることが多いため、除外した。この結果、起点となるストーリー S_0 として 1,431 ストーリーが抽出された。次に、各 S_0 を起点とするトピックスレッド構造 T_r を構築した。その結果、 S_0 を含む 2 ストーリー以上からなるトピックスレッド構造 437 件が得られた。なお、各トピックスレッド構造を得るために、Sun Blade 1000 ワークステーションを用いて、構造の複雑さによって数秒から数 10 分の処理時間を要した。

4.2 トピックスレッド構造の内容に関する分析

得られた構造のなかには、図 2 のようにトピックの流れを把握するうえで役立ちそうな構造だけでなく、解釈が困難なものもあった。これは設定した閾値の加減によると思われるが、一方で極めて小規模な構造も一定数あり、十分トピックの流れを追跡できていないものと思われる。また、ストーリー分割の失敗により、本来無関係と思われるトピックが混入したものもあった。今後は、得ら

*2 理想的にはアーカイブ全体を探索したいが、計算時間を抑制するため、探索期間を限定した。

表 3 トピックスレッド構造の規模.

	最小	平均	最大
ストーリー数	2	15.7	85
映像長 [秒]	5	4,380	38,125
(総映像長比)	(0.0029%)	(2.6%)	(22%)

れる構造の規模や複雑さに応じた適切な閾値の設定方法や利用者の操作による適合フィードバックの導入, より高精度なストーリー分割手法の導入を検討する必要がある.

得られた構造のなかには, 図 2 や図 5 に例示したように, 記憶や実体験と照合して, ニューストピックの変遷を無理なく理解できるものもあるが, 明らかに実社会において無関係なストーリーからなるトピックスレッド構造も得られた.

これには, 以下の 2 つの原因がある.

- ストーリー分割の失敗

3・2 節の手順によるストーリー分割の失敗により, 本来複数のストーリーに分割されるべきものが 1 つのストーリーになった場合, 無関係なトピックに関連付いてしまう.

- ストーリー間の関連性評価の問題

3・3 節の手順 (1) において評価した関連性が実社会における関連性を正しく反映していないことがあった. たとえば, 日常的に発生する類似した事件や気象現象に関する報道などは現状のキーワードベクトルの比較方法では, 誤って関連すると判断されやすい.

以上の点をふまえ, 今後ストーリー分割性能の向上と, より正確な関連性評価方法を検討する. 前者は, 画像特徴を手がかりとした手法と組み合わせることが考えられる. 後者は, 関連ストーリーの探索範囲 (期間) を制限することや, ストーリーのジャンル毎に特定の属性をもつキーワードに重みを与えることなどを考えている.

4・3 トピックスレッド構造に基づく閲覧の効率化

次に, トピックスレッド構造を用いることで, どのように閲覧が効率化され得るのかについて, 探索範囲の全映像と構造中に含まれる映像長を比較した. なお, ここでは構造の妥当性は判断せずに, 提案手法で関連があると判断された映像のみを閲覧することで, 何の手がかりもなくその期間の全映像を閲覧して取捨選択するのと比較して, どの程度閲覧時間が短縮されるかを評価尺度とする.

§1 トピックスレッド構造の規模

抽出された 437 件のトピックスレッド構造の規模について分析する. 表 3 に, トピックスレッド構造あたりに含まれるストーリー数及びそれらに対応する映像長を示す.

関係するストーリー数が少なければ, ある程度人手で整理することもできるが, 実際には 5 ストーリー以上を含む規模が大きいトピックスレッド構造が半数以上あり, トピックスレッド構造を自動構築する必要性がうかがえる.

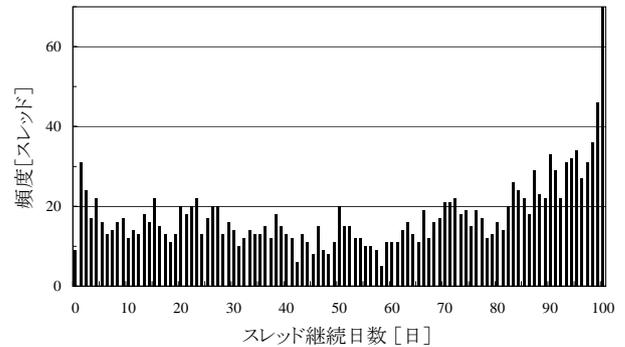


図 6 トピックスレッド継続日数の分布.

表 4 ストーリー単位で見るトピックスレッドの規模.

	最小	平均	最大
ストーリー数	2	7.68	29
映像長 [秒]	5	2,770	19,038
(対総映像長)	(0.0029%)	(1.6%)	(11%)

また, トピックスレッド構造に含まれる映像長は, あるストーリーから派生する全てのトピックスレッドに含まれるストーリーの映像を閲覧するのに要する時間である. 探索期間 100 日の総映像長 47.7 時間に対し, トピックスレッド構造を用いることにより, 最悪でも 22% の約 10.6 時間, 平均では 2.6% の 1.22 時間の映像を閲覧すれば派生的な話題を含めて網羅的にアーカイブ中に存在する関連映像の内容を把握できる.

§2 トピックスレッドの規模

一般に, 利用者がトピックスレッド構造全体を理解する必要があることは少なく, 実際には感心があるいくつかのトピックスレッド (起点と葉をつなぐ任意の経路) 上の映像を順に閲覧すれば満足すると考えられる. そこで, 上記 437 件のトピックスレッド構造に含まれるトピックスレッド 1,788 本について, その規模を分析する.

まず, 図 6 に, トピックスレッドあたりの継続日数 (起点から葉までの日数) の分布を示す. 平均は 56.4 日 / トピックスレッドだが, 0 日^{*4}から 100 日までまんべんなく分布している. 継続期間 100 日に大きな値が現われることは, 探索期間を延ばせば, より長く続き得るトピックスレッドが一定数あることを示唆している.

次に表 4 に, トピックスレッドあたりに含まれるストーリー数及びそれらに対応する映像長の総和を示す. 探索期間 100 日の総映像長 47.7 時間に対し, 利用者の関心がある話題の流れに沿ったトピックスレッドを適切に選択できれば, 最悪でも 11% の約 5.3 時間, 平均では 1.6% の 46 分の映像を閲覧すれば, アーカイブ中に存在する関連映像の内容を把握できる.

*4 起点と同一日の後続するストーリーで完結する場合.

表 5 トピッククラスタ単位で見るトピックスレッドの規模 .

	最小	平均	最大
トピック			
クラスタ数	1	3.08	9
映像長 [秒]	5	1,280	6,741
(対総映像長)	(0.0029%)	(0.75%)	(3.9%)

§ 3 トピッククラスタの効果

トピックスレッドを適切に選択できることによる絞込み効果は大きいですが、4・3・2 節で得られた平均的な映像長は依然比較的長時間である。そこで、トピックスレッド上で類似した隣接ストーリーをまとめたトピッククラスタを用いて更に効率的に映像を閲覧することが考えられる。ここでは、各トピックスレッドにおけるトピッククラスタ構造の規模について分析し、その効果を分析する。

表 5 に、スレッドあたりに含まれるクラスタ数及び各トピッククラスタの先頭ストーリーに対応する映像長の総和の分布を示す。このなかには、全 17 ストーリーが 1 トピッククラスタに属するもの、28 ストーリーが 4 トピッククラスタに分かれて属するものなどがあり、各トピッククラスタの規模はまちまちだった。

探索期間 100 日の総映像長 47.7 時間に対し、利用者の関心がある話題の流れに沿ったトピックスレッドを適切に選択し、各トピッククラスタの先頭ストーリーのみを閲覧すれば、最悪でも 3.9% の約 1.9 時間、平均では 0.75% の約 21 分の映像を閲覧すれば、アーカイブ中に存在する関連映像の内容を大雑把に把握できる。また、この際にトピックスレッド 1 本単位での映像長の平均短縮率は 46% だった。

5. 映像閲覧・編纂支援インターフェース

以上の手法に基づいて得られる任意のストーリーを起点としたトピックスレッド構造を可視化し、ニュース映像アーカイブ中の膨大な映像を時間的な変遷を追跡しながら閲覧・映像編纂を支援するインターフェース: mediaWalker を試作した。以下にその機能を紹介し、関連研究と比較する。

5.1 閲覧・追跡機能

mediaWalker では、まず図 7 (a) にあるようなインターフェースにより、利用者の関心があるストーリー (起点となるストーリー) を指定する。ここで一覧表示されるストーリーは、何らかの手段により検索された結果や、指定された日に放送されたストーリーの一覧などである。利用者は各ストーリーを試聴しながら、どれを起点とするか選択する。

同図 (b) には、同図 (a) でサムネイル画像の右側のボタンを押して、そのストーリーを起点とする未来方向のトピックスレッド構造が、トピッククラスタ構造と重畳して表示される様子を示している。



図 7 トピックスレッド構造に基づくニュース映像閲覧・追跡インターフェース: mediaWalker . 図 5 と同じ構造を表示。サムネイル画像はストーリーに、背景の塗りつぶしはトピッククラスタに対応。

各ストーリーはサムネイル画像で表示され、各々を選択すると映像が拡大再生されるとともに、画面左下部にそのストーリーの詳細情報が表示され、放送日時やキーワード一覧などを見ることが出来る。これと同時に、図 8 に示すように、選択したストーリーの前後のストーリーとの間のエッジが強調され、それらのストーリーのサムネイルが画面内にちょうど収まるように自動的に表示領域が拡大縮小される。また、エッジ上にマウスを移動することで、エッジ前後のストーリーにおけるキーワードの変化が表示される (図 8 (b) に例示) 。

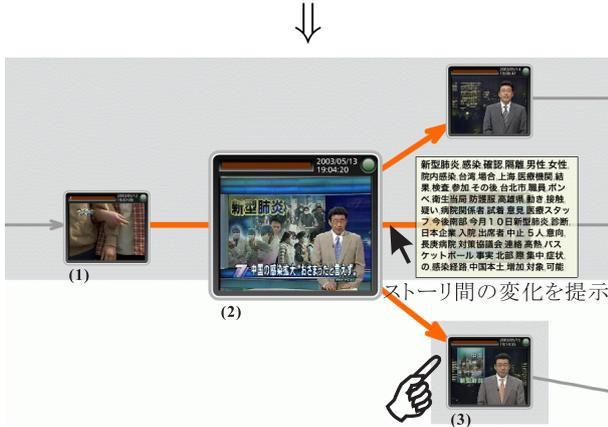
これらの機能により、前後の話題の流れを把握しながら、順を追ってトピックスレッド上の映像を閲覧することができる。一方、図 7 (a) でサムネイル画像の左側のボタンを押すことで、過去方向への流れを見ることが出来る。これにより、関心があるストーリーがどのような経緯により生じたかを閲覧することも出来る。

5.2 映像編纂支援機能

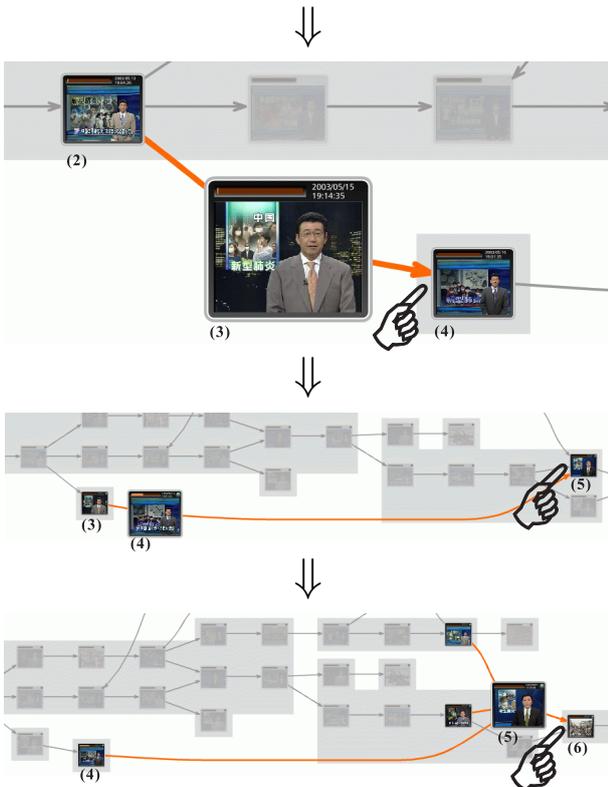
一方、映像編纂を支援するために、上記一連の機能に加えて、ストーリーやトピックスレッドを選択し、プレイリストとして出力する機能を備えた。図 9 に示すように、プレイリストに含めたいストーリーについて、サムネイル画像右上の四角にチェックを入れることで出力し、選択されたストーリーは強調される。また、指定したストーリーを含むトピックスレッド、つまり指定したストーリー間を説明し得る一連の映像の集合を自動的に抽出して出力する機能も備えた。同様に、選択されたスレッドは緑色で強調される。なお、指定されたストーリーを含む複数のト



(a) 注目しているストーリーの映像を拡大再生し、前後のストーリーとの間のエッジを橙色に強調。前後のストーリーが画面内に収まるよう、自動的に拡大縮小表示。次のストーリーを選択することで、注目ストーリーを変更。



(b) エッジ上にマウスを重ねることで前後のストーリーとの間のキーワードの変化を提示し、特に分岐点において追跡したいスレッドの選択を支援。



(c) このように次々とストーリーを閲覧・追跡することで、関心があるトピックの流れの理解を支援。

図 8 図 7 (b) 中のトピックスレッド構造中のストーリーを順にたどって閲覧する様子。



図 9 プレイリスト機能：図 8 中のストーリー (1) (4) (6) を結ぶ経路を自動抽出し、プレイリストを出力した例。

ピックスレッドが存在する場合は、候補が薄く強調され、利用者に一意に確定するよう促す。

以上のようにしてプレイリストが出力されれば、それを素材映像として既存の映像編集インタフェースや、自動要約映像作成手法を利用するなど、様々な形で再編集することができる。大規模映像アーカイブにおいては、利用者が求める流れに沿った素材映像を抽出することすら困難であるため、前節までに示した数値からも、提案手法及びインタフェースを利用することで、アーカイブ中の映像を効率的に利用できるようになると考えられる。

5.3 関連研究との比較

2.2 節で紹介した一連の研究と本インタフェースを比較する。

まず、本インタフェースでは、各映像を代表する画像として、先頭フレームを表示した。しかし今後は、関連研究で紹介したような手法により、よりよく映像内容を代表するフレームを選択することを考えたい。

次に、ニュース映像の時系列の流れを可視化・閲覧するインタフェースに関して、関連研究では、起点となる映像から 1 列に連なった映像を順に閲覧し、場合によって相互に比較するようになっている。これに対し、本インタフェースでは、分岐・合流する現実世界のトピックの流れが有向グラフとして表現されるため、利用者が関心をもっている流れに絞り込んで効率的に閲覧できるほか、構造全体が可視化されることで、トピックの広がりを直感的に把握したり、複雑に絡み合うトピック間の因果関係を把握しやすくなると期待される。

一方、放送アーカイブ中の映像の再編集に関する研究としては、Casaresらによる Silver インタフェース [Casares 02] があるが、これは特定の映像を再編集するためのものであり、本インタフェースのように、関連する複数の映像を再編集するためのものではない。

6. 結 論

本論文では、まずニュースストーリー間の時系列意味構造：トピックスレッド構造を提案し、その構造を用いることで、膨大な量の映像に対してどの程度閲覧が効率化

され得るかを分析した。

次に、トピックスレッド構造を可視化し、関心があるトピックスレッドを選択しながら閲覧してだけで、利用者がトピックの流れを理解し、それを説明する映像を再編集することを支援するインタフェースを紹介した。

今後の課題としては、解析処理の各段階における要素技術の高速・高精度化のほか、スレッド構造中の各要素（ストーリー、トピッククラスタ、トピックスレッド）を特徴付ける情報の抽出と可視化、インタフェースにおいてスレッドを選択しやすい情報の提示、利用者の選択に基づく適応フィードバックによる動的なトピックスレッド再構築手法の検討などが挙げられる。また、得られた構造を用いた自動映像要約 [Ide 06] や、新聞記事など他メディアへの適用も検討している。

謝 辞

本研究の一部は、文部科学省科学研究費補助金（特定領域研究「情報学」:15017285, 16016289, 同「情報爆発IT基盤」:18049035, 19024036, 若手研究(B):15700116, 18700080), 栢森情報科学振興財団一般研究助成(K17研X202)及び国立情報学研究所における名古屋大学及び(株)ネットコンパスとの共同研究(株)三菱総合研究所からの受託調査による。また、インタフェースの実装に貢献してくださった(株)ネットコンパスの菟川武治氏に感謝する。

◇ 参 考 文 献 ◇

- [Casares 02] Casares, J., Myers, B., Long, A., Bhatnagar, R., Stevens, S., Dabbish, L., and Corbett, A.: Simplifying video editing with intelligent interaction, in *Adjunct Proc. CHI2002: Human Factors in Computing Systems*, pp. 672–673 (2002)
- [Christel 00] Christel, M., Olligschlaeger, A., and Huang, C.: Interactive digital maps for a digital video library, *IEEE Multimedia*, Vol. 7, No. 1, pp. 60–67 (2000)
- [Christel 02] Christel, M., Hauptmann, A., Wactler, H., and Ng, T.: Collages as dynamic summaries for news video, in *Proc. 10th ACM Int. Conf. on Multimedia*, pp. 561–569 (2002)
- [Duygulu 04] Duygulu, P., Pan, J.-Y., and Forsyth, D.: Towards auto-documentary: Tracking the evolution of news stories, in *Proc. 12th ACM Int. Conf. on Multimedia*, pp. 820–827 (2004)
- [Ide 05] Ide, I., Kinoshita, T., Mo, H., Katayama, N., and Satoh, S.: trackThem: Exploring a large-scale news video archive by tracking human relations, in Lee, G., Yamada, A., Meng, H., and Myaeng, S.-H. eds., *Information Retrieval Technology, 2nd Asia Information Retrieval Symposium, AIRS2005, Jeju Island, Korea, October 2005, Procs.*, Vol. 3689 of *Lecture Notes in Computer Science*, pp. 510–515, Springer-Verlag, Berlin, Heidelberg (2005)
- [Ide 06] Ide, I., Mo, H., Katayama, N., and Satoh, S.: Exploiting topic thread structures in a news video archive for the semi-automatic generation of video summaries, in *Proc. 2006 IEEE Int. Conf. on Multimedia and Expo*, pp. 1473–1476 (2006)
- [Katayama 04] Katayama, N., Mo, H., Ide, I., and Satoh, S.: Mining large-scale broadcast video archives towards inter-video structuring, in Aizawa, K., Nakamura, Y., and

- Satoh, S. eds., *Advances in Multimedia Information Processing —PCM2004 5th Pacific Rim Conf. on Multimedia, Tokyo, Japan, November/December 2004, Procs. Part II—*, Vol. 3332 of *Lecture Notes in Computer Science*, pp. 489–496, Springer-Verlag, Berlin, Heidelberg (2004)
- [Nallapati 04] Nallapati, R., Feng, A., Peng, F., and Allan, J.: Event threading within news topics, in *Proc. 13th ACM Conf. on Information and Knowledge Management*, pp. 446–453 (2004)
- [Nat 00] National Institute of Standards and Technologies: *The year 2000 Topic Detection and Tracking (TDT2000) task definition and evaluation plan* (2000), <http://www.nist.gov/speech/tests/tdt/tdt2000/evalplan.htm>
- [Rautiainen 04] Rautiainen, M., Ojala, T., and Seppanen, T.: Cluster-temporal browsing of large news video databases, in *Proc. 2004 IEEE Int. Conf. on Multimedia and Expo*, Vol. 2, pp. 751–754 (2004)
- [Snoek 05] Snoek, C., Worring, M., van Gemert, J., Geusebroek, J.-M., Koelma, D., Nguyen, G., de Rooij, O., and Seinstra, F.: MediaMill: Exploring news video archives based on learned semantics, in *Proc. 13th ACM Int. Conf. on Multimedia*, pp. 225–226 (2005)
- [Uramoto 98] Uramoto, N. and Takeda, K.: A method for relating multiple newspaper articles by using graphs, and its application to webcasting, in *Proc. 36th Annual Meeting of the Association for Computational Linguistics*, Vol. 2, pp. 1307–1313 (1998)
- [Wu 06] Wu, X., Ngo, C.-W., and Li, Q.: Threading and autodocumenting news videos, *IEEE Signal Processing Mag.*, Vol. 23, No. 2, pp. 59–68 (2006)
- [Yang 06] Yang, C., Shi, X., and Wei, C.-P.: Tracing the event evolution of terror attacks from on-line news, in Chen, H., Wang, F., Yang, C., Zeng, D., and Chau, M. eds., *Intelligent and Security Informatics —IEEE Int. Conf. on Intelligent and Security Informatics, ISI 2006, San Diego, CA, USA, May 23–24, 2006, Procs.* —, Vol. 3975 of *Lecture Notes in Computer Science*, pp. 343–354, Springer-Verlag, Berlin, Heidelberg (2006)
- [Zhai 05] Zhai, Y., Yilmaz, A., and Shah, M.: Story segmentation in news videos using visual and text cues, in Leow, W.-H., Lew, M., Chua, T.-S., Ma, W.-Y., Chaisorn, L., and Bakker, E. eds., *Image and Video Retrieval —4th Int. Conf., CIVR2005, Singapore, July 20–22, 2005, Procs.* —, Vol. 3568 of *Lecture Notes in Computer Science*, pp. 92–102, Springer-Verlag, Berlin, Heidelberg (2005)
- [井手 02] 井手 一郎, 浜田 玲子, 坂井 修一, 田中 英彦: テレビニュース字幕の語義属性解析のための辞書作成, 電子情報通信学会論文誌(D-II), Vol. J85-D-II, No. 7, pp. 1201–1210 (2002)
- [井手 03] 井手 一郎, 孟 洋, 片山 紀生, 佐藤 真一: 大規模ニュース映像コーパスの意味構造解析, Technical Report PRMU2003-97, 電子情報通信学会パターン認識・メディア理解研究会 (2003)
- [谷口 99] 谷口 行信, 阿久津 明人, 外村 佳伸: PanoramaExcerpts: パノラマ画像の自動生成・レイアウトによる映像一覧, 電子情報通信学会論文誌(D-II), Vol. J82-D-II, No. 3, pp. 390–398 (1999)

〔担当委員：榊井 文人〕

2007年11月30日 受理

 著 者 紹 介



井手 一朗(正会員)

1994 年 東京大学工学部電子工学科卒業。2000 年 同大学大学院工学系研究科電気工学専攻博士課程修了。博士(工学)。同年 国立情報学研究所助手。2004 年 名古屋大学大学院情報科学研究科助教授, 国立情報学研究所客員助教授兼任。2007 年 同准教授, 現在に至る。電子情報通信学会, 情報処理学会, IEEE Computer Society, ACM 各会員。



木下 智義(正会員)

1995 年 東京大学工学部電子情報工学科卒業。2000 年 同大学大学院工学系研究科情報工学専攻博士課程修了。博士(工学)。現在 (株) ネットコンパス代表取締役社長。電子情報通信学会, 情報処理学会各会員。



高橋 友和

1997 年 茨城大学工学部情報工学科卒業。2003 年 同大学大学院理工学研究科博士後期課程修了。博士(工学)。同年 名古屋大学大学院情報科学研究科 21 世紀 COE 研究員。2005 年 日本学術振興会特別研究員(PD)。2008 年 岐阜聖徳学園大学経済情報学部准教授, 現在に至る。電子情報通信学会, 画像電子学会各会員。



孟 洋

1990 年 武蔵工業大学工学部電気電子工学科卒業。1997 年 東京大学大学院工学系研究科電子工学専攻博士課程修了。博士(工学)。同年 日本学術振興会博士特別研究員。1998 年 学術情報センター助手。2000 年 国立情報学研究所助手。2007 年 同助教, 現在に至る。電子情報通信学会, 情報処理学会, 日本ファジィ学会, 映像情報メディア学会各会員。



片山 紀生

1990 年 東京大学工学部電気工学科卒業。1995 年 同大学大学院工学系研究科電気工学専攻博士課程修了。博士(工学)。同年 学術情報センター助手。2000 年 国立情報学研究所助教授。2007 年 同准教授, 現在に至る。電子情報通信学会, 情報処理学会, IEEE, ACM 各会員。



佐藤 真一

1987 年 東京大学工学部電子工学科卒業。1992 年 同大学大学院工学系研究科情報工学専攻博士課程修了。工学博士。同年 学術情報センター助手。1998 年 同助教。2000 年 国立情報学研究所助教授。2004 年 同教授, 現在に至る。電子情報通信学会, 情報処理学会, IEEE Computer Society, ACM 各会員。



村瀬 洋

1978 年 名古屋大学工学部電気電子工学科卒業。1980 年 同大学大学院電気電子工学専攻修士課程修了。同年 日本電信電話公社(現 NTT)入社。1992 年より 1 年間米国コロンビア大客員研究員。2003 年 名古屋大学大学院情報科学研究科教授, 現在に至る。工学博士。電子情報通信学会, IEEE 各フェロー, 情報処理学会各会員。