

意味構造に基づく大規模ニュース映像アーカイブの検索と閲覧

井手 一郎 (名古屋大学/国立情報学研究所)

Retrieval and browsing in a large-scale news video archive according to its semantic structure

Ichiro IDE (Nagoya University / National Institute of Informatics)

1. まえがき

記憶装置の大容量化を活かし、我々は2001年3月より今日にいたるまで、特定のニュース番組(NHK News7)の映像を連日蓄積してきた [1]。現在、およそ1,900日分、のべ900時間の映像と、付随して放送された文字放送字幕(音声書き下しテキスト)を蓄積している。

従来、一般の情報検索同様、ニュース映像に対しても、キーワード入力による検索手法が提案されてきた。しかし、アーカイブの規模が大きくなるにつれ、大量の検索結果を1つ1つ閲覧しながら所望の映像を探す作業の負担が著しく大きくなる。

そこで我々は、ニュースストーリー間の意味関係に基づく時系列的なトピック構造に着目し、意味内容に沿った検索・閲覧を通じて、効率的な映像内容理解を支援することを目指している。

本講演では、このような目的を実現するための意味構造(トピックスレッド構造)解析手法について簡単に述べた後、得られた構造に基づく検索・閲覧インタフェースを紹介する。

2. ニュース映像アーカイブからの意味構造抽出

2.1 ニュース映像アーカイブにおける意味構造

図1に示すように、1本のニュース映像は複数のストーリーからなる。これを映像内の意味構造と考える。一方、実世界における事態の推移に伴い、複数日にまたがるストーリーの集合がトピックを形成する。これを映像間の意味構造と考える。

本研究では、最終的に映像間の意味構造を抽出するが、そのためにはまず、映像内の意味構造を解析する必要がある。以下の2節で各々の処理について簡単に述べる。なお、手法の詳細については、文献[2]を参照されたい。

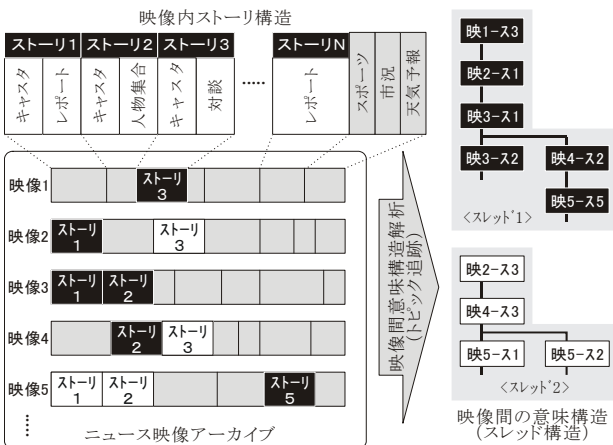


図1: ニュース映像アーカイブにおける映像内・映像間の意味構造

2.2 映像内構造の解析: ストーリー分割

映像内の意味構造、つまり図1左側に示すようなストーリー構造の抽出に関する研究はこれまでも多くなされてきた。その多くは、「キャストショット」と呼ばれる特徴的な映像の出現を手がかりとしたものだったが、番組毎のデザインに依存する問題があった。

そこで、我々は映像に付随して放送される文字放送字幕と呼ばれる、音声を書き下したテキストに対する自然言語処理を手がかりにする手法を採ることとした。手法の概要を以下に示す。

- 各文に形態素解析を施し、名詞列を抽出する。
- 抽出された名詞列から、出現頻度付きキーワードベクトルを作成する。
- ある文の境界において、前後 w 文のキーワードベクトルの類似度を余弦尺度で評価する。
- 各窓幅 w における類似度の最大値が閾値 θ_{seg} を下回れば、その箇所にストーリーの境界を検出する。

以上の処理により、両端で前後1文の誤差を許せば、再現率95.4%、適合率90.5%の精度でストーリー境界を検出できることが、約1月分のニューステキストに対する実験により確認された。

2.3 映像間構造の解析: トピックスレッド構造の抽出

次に、得られたストーリー間の意味構造を解析する。

まず、全てのストーリー間の類似度を事前に評価しておく。類似度は2.2で用いたのと同様に、ストーリー内の名詞列をキーワードとした出現頻度付きキーワードベクトルの余弦尺度で評価する。

ここまでの処理は全て自動化されており、毎晩の放送後、翌朝までには最新の状態に更新が完了している。

次に、類似度が強いストーリー対を時系列に連結していくことで、スレッド構造を得る。スレッド構造の抽出手順の概要は以下の通りである。

- 注目ストーリーを根とし、類似度が強いストーリーを再帰的に展開した木を作る。兄弟間の並び方は、ストーリーの時系列順とする。
- 木構造中に存在する、同一ストーリーを親とする部分木につき、最も左のもの1つに統合する。この際、リンクも統合するため、最終的に得られる構造は有向グラフになる。

この処理を再帰的に施すことにより、意味的なつながりのあるストーリーを時系列順につないだスレッド構造が得られる。結果として、トピックの分岐や合流などの流れを理解する手がかりになる。この処理は計算量が膨大であるため、必要に応じてその都

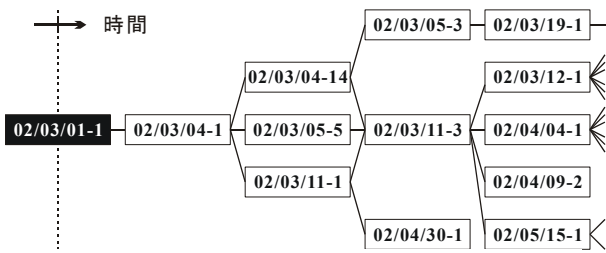


図 2: 注目ストーリーを起点としたトピックスレッド構造

度行われる。また、現実的な処理時間で結果を得るために、探索範囲(日数・階層など)を限定する。以上の処理により、実際に得られたスレッド構造の例を図 2 に示す。

3. 意味構造に基づく検索・閲覧インターフェース

以下に、得られたスレッド構造を用いたニュース映像の検索・閲覧インターフェースを 2 つ紹介する。

3.1 スレッド構造に基づくトピック追跡インターフェース

本インターフェースは、スレッド構造上のストーリーを 1 つ 1 つ選択しながらどって行くものである。図 3 にスナップショットを示す。

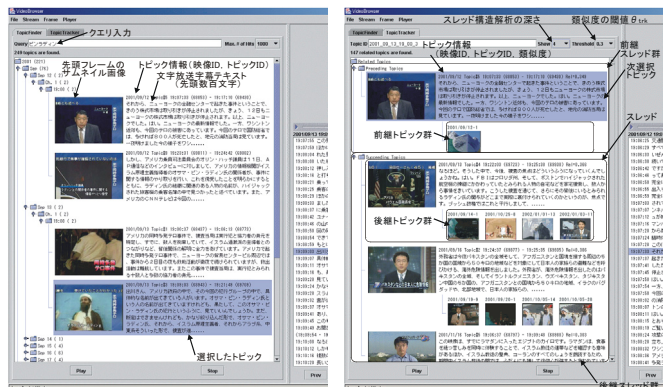
まず、図 3(a)に示す topicFinder と呼ばれる従来のキーワード検索インターフェースにより、始点となるストーリーを指定する。次に、インターフェース中の「track」ボタンにより、図 3(b)に示す topicTracker インターフェースに切り替える。このインターフェースで表示されるのは、スレッド構造における直前・直後のストーリーである。この中から興味があるものを閲覧し、繰り返し追跡していくことで、スレッド構造をたどっていく。スレッド構造は、各ノードにおける分岐を極力少なくなるように設計されているため、興味あるトピックの流れを効率的に追えることになる。

3.2 スレッド構造を可視化したインターフェース

3.1 で紹介した topicTracker は、トピックの追跡には適していたが、トピックの大きな流れやスレッド間の差を把握しにくかった。そこで図 4 に示すような、スレッド構造そのものを可視化した threadViewer インターフェースを作成した。このインターフェースにより、より大局的な意味構造の理解が可能になると期待される。

4. おわりに

本講演では、大規模なニュース映像アーカイブにおけるトピックスレッド構造を抽出し、効率的な検索・閲覧を実現する手法を紹介した。現在はテキスト情報のみを利用しているが、現在、同一(near-duplicate)映像区間の高速検出手法[4]を開発しているため、今後は画像の共起性も利用した意味構造の解析を実現していく。また、得られた意味構造のより効率的な内容理解を実現するために、スレッド構造に基づく映像要約についても検討している[3]。



(a) topicFinder インターフェース (b) topicTracker インターフェース

図 3: スレッド構造に基づくニューストピック追跡インターフェース

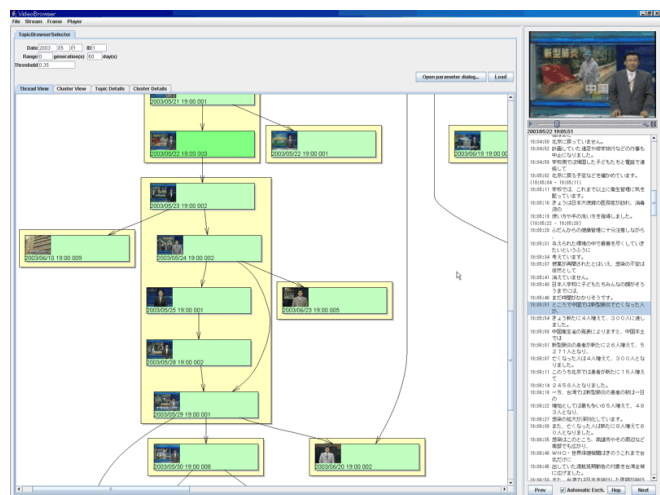


図 4: スレッド構造を可視化した threadViewer インターフェース

謝辞

本講演の内容は村瀬洋, 高橋友和, 学生諸氏(名古屋大学), 佐藤真一, 片山紀生, 孟洋(国立情報学研究所)らとの共同研究の成果である。また本研究の遂行にあたり、文部科学省・日本学術振興会より 21 世紀 COE プログラム及び科学研究費補助金を、栢森情報科学振興財団より研究助成を受けている。

文献

- (1) N. Katayama, H. Mo, I. Ide, S. Satoh: "Mining large-scale broadcast video archives towards inter-video structuring," 5th Pacific Rim Conf. on Multimedia, Dec. 2004.
- (2) 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," 信学技報, PRMU2003-97, Sept. 2003.
- (3) I. Ide, H. Mo, N. Katayama, S. Satoh: "Exploiting topic thread structures in a news video archive for the semi-automatic generation of video summaries," 2006 IEEE Intl. Conf. on Multimedia and Expo, July 2006.
- (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.