

クローズドキャプションを用いた ニュース番組におけるトピック分割手法

Topic Segmentation in News Programs Based on Closed Captions

樋上 直樹† 片岡 充照† 黒木 修隆† 沼 昌宏† 山本 啓輔††
Naoki Hinoue Mitsuteru Kataoka Nobutaka Kuroki Masahiro Numa Keisuke Yamamoto

1. まえがき

近年、放送形態の多様化や大容量 HDD レコーダの普及によって視聴者が蓄積する映像情報は増加している。このような膨大な映像の整理・視聴を支援するため、番組の解析・構造化の研究が行われている。ニュース番組については、番組内のトピックの分割や、番組間のトピックの収集・整理が必要とされている。映像処理や音声処理を伴う従来のトピック分割手法 [1] は、民生機用の MPU にとって大きな負荷となることが問題である。本研究では近年付与率が向上しつつあるクローズドキャプションに注目し、テキスト解析によってトピック分割を行う。新たに定義する文章結合力に基づいて、分割処理をリアルタイムに進める点を特徴とする。実験では実際に放送されたニュース番組を使用し、省メモリ・高精度なトピック分割が可能であることを示す。

2. トピック分割手法

2.1 クローズドキャプション

クローズドキャプションは、デジタル放送において映像と連動して得られるテキスト情報である。ニュース番組におけるクローズドキャプションは、アナウンサの発話をほぼそのまま表示しており、これを読むだけでもニュース内容を理解することができる。クローズドキャプションの表示例を図 1 に示す。ニュース番組においては、トピックごとに出現する単語の傾向が大きく変化するという特徴がある。以後、第 i 番目に出現する文を S_i とする。



図 1 クローズドキャプション表示例

2.2 キーワードの抽出

従来のテキストセグメンテーションでは、テキストをデータベースに照合しながらキーワードを抽出するが、データベースにメモリを圧迫される問題がある。提案手法ではデータベースを用いず、図 2 のように文字種でテキストを区切る。そして「ひらがな」と「記号」以外の文字種が 2 文字以上連続するものをキーワードとする。文 S_i に含まれるキーワードの頻度をキーワードベクトル K_{S_i} で表す。

2.3 従来の文間類似度の定義

Hearst のテキストセグメンテーション [2] では、ある境界から前後一定範囲の窓内で単語の出現頻度によるベクトルを生成し、両者の内積によって単語の結束度（文間の類似度）を求める。境界 i における単語の結束度 $C(i)$ は、前述のキーワードベクトル K_{S_i} を用いると、

$$C(i) = \frac{\sum_{m=i-w+1}^i K_{S_m} \cdot \sum_{n=i+1}^{i+w} K_{S_n}}{\left| \sum_{m=i-w+1}^i K_{S_m} \right| \cdot \left| \sum_{n=i+1}^{i+w} K_{S_n} \right|} \quad (1)$$

と表される。松井 [3] や孟 [4], [5] とも同様の計算に基づいて文間の類似度を求めている。この $C(i)$ が極小となる点がトピック分割の候補となる。

2.4 提案する文章結合力の定義

提案手法では文 S_i について、それより過去の文への結合力 $C_P(i)$ と、未来の文への結合力 $C_F(i)$ を定義する。文章結

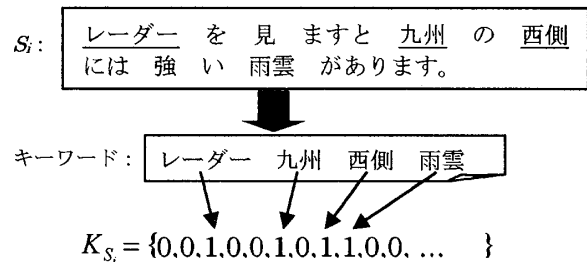


図 2 キーワード抽出例

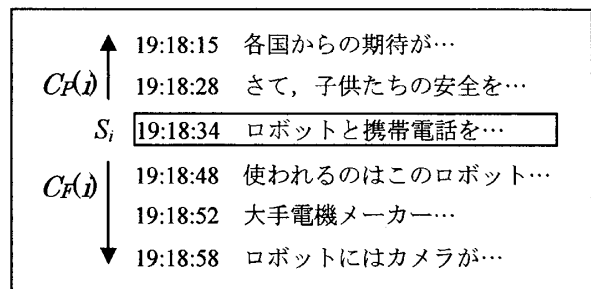


図 3 文章結合力の概念

† 神戸大学大学院工学研究科
Graduate School of Engineering, Kobe University.
†† 神戸大学連携創造本部
Headquarters for Innovative Cooperation and
Development, Kobe University.

合力の概念を図3に示す。これらは前述のキーワードベクトル K_{S_i} を用いて、

$$C_P(i) = \sum_{x=1}^X K_{S_i} \cdot K_{S_{i-x}} \quad (2)$$

$$C_F(i) = \sum_{y=1}^Y K_{S_i} \cdot K_{S_{i+y}} \quad (3)$$

のように求める。ただし、 X, Y は S_i から見て比較時間 T 秒以内に存在する文の数である。

2.5 表層情報を用いた補正

文頭に転換の接続詞が存在すればトピックが分割されるよう結合力に補正を加える。逆に、文頭に指示代名詞や順接などの接続詞が存在すればトピックが分割されないよう結合力に補正を加える。

2.6 トピックの分割

ある文 S_i について、過去との結合力が閾値 T_h より強く、未来との結合力が T_h より弱いとき、すなわち、

$$C_P(i) > T_h \geq C_F(i) \quad (4)$$

を満たすとき、 S_i はトピックの末尾文である可能性が高い。逆に文 S_j が、

$$C_P(j) \leq T_h < C_F(j) \quad (5)$$

を満たすとき、 S_j はトピックの先頭文である可能性が高い。そこでこれらの文が $i < j$ の順に現れたとき、その間でトピックの転換があったと考え分割を行う。

放送中のクローズドキャプションへの適用を考えた場合、 $C_P(i)$ は未来の文の受信を待たずに値を確定できるが、 $C_F(i)$ は、 T 秒以内に新たな文が出現する度に値を再計算しなければならない。

2.7 映像へのリンク

分割したトピックに、そのトピックにおける先頭の一文と、トピック内の頻出単語を合わせたものをタイトルとして付与し、映像と対応付ける。アプリケーションの一例を図4に示す。これによりユーザは自分の見たいトピックに素早く到達することができる。

3. 実験と考察

NHKのニュース7(19日分)のクローズドキャプションに対して従来手法[2]および提案手法を適用し、分割精度および処理時間を比較する。ただし、従来手法における窓幅 $w=8$ とし、提案手法における比較時間 $T=150(\text{sec})$ 、閾値 $T_h=3$ とする。

3.1 評価方法

分割精度の指標としては正確性を表す適合率、網羅性を表す再現率と、それらの調和平均であるF値を用いる。また処理時間としては、番組が終了した後、解析が終了するまでの時間を比較する。

3.2 実験結果および考察

実験結果を表1に示す。分割精度について、提案手法は従来手法をF値で22.7ポイント上回った。また、処理時間について、提案手法は従来手法の35%程度に抑えることができた。この要因は、提案する文章結合力を用いることにより演算量を削減でき、かつニュース番組に適したアルゴリズムを構築できたためと考えられる。

表1 分割精度および処理時間

手法	従来	提案
適合率(%)	64.1	80.8
再現率(%)	51.6	79.2
F値(%)	56.9	79.6
処理時間(ms)	1112	385

4. まとめ

本研究ではテレビ番組の視聴支援のため、クローズドキャプションを利用し、従来より省メモリ・高精度かつ放送中の番組を解析可能なニューストピック分割手法を提案した。今後の課題として、分割したトピックの収集と整理が挙げられる。

参考文献

- [1] 井手一郎, 山本晃司, 浜田玲子, 田中英彦, “ショット分類に基づく映像への自動的索引付け手法”, 電子情報通信学会論文誌, Vol.J82-D-II, No.10, pp.1543-1551, 1999年10月.
- [2] M.A.Hearst, “TextTiling: Segmenting Text into Multi-paragraph Subtopic Passages”, Computational Linguistics, Vol.23, No.1, pp.33-64, 1997.
- [3] 松井祥峰, 乾伸雄, 小谷善行, “単語の結束度と文の表層情報を組み合わせたテキストセグメンテーション”, 情報処理学会研究報告, 2004-NL-162, Vol.2004 No.73, pp.151-158, 2004年7月.
- [4] 佐藤真一, 孟洋, 加藤弘之, 井手一郎, “複数メディアによる大規模ニュース映像群の意味的構造解析”, 国立情報学研究所, 研究成果報告書, A02-27, 2004年3月.
- [5] 孟洋, 山岸史典, 井手一郎, 佐藤真一, 坂内正夫, “ニュース映像アーカイブにおけるキーショットの抽出と索引付け”, 信学技報, PRMU2005-31, 2005年6月.

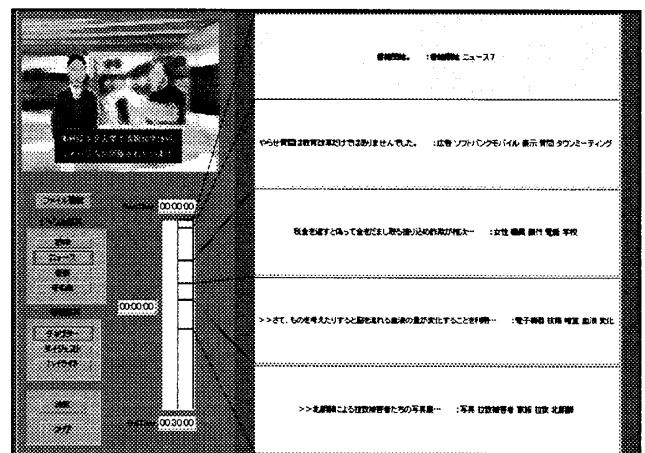


図4 視聴支援アプリケーション例