

クローズドキャプションを利用した番組紹介コメント 自動生成の検討

A Study on Automatic Generation of Commentaries for a TV Program Trailer Using Closed Captions.

小林 彰太† 山田 一郎†† 柴田 正啓†† 八木 伸行†† 羽鳥 光俊†
Shota Kobayashi Ichiro Yamada Shibata Masahiro Nobuyuki Yagi Mitutoshi Hatori

1. はじめに

近年、家庭のハードディスクレコーダーの容量が増加し、大量のテレビ番組を蓄積できるようになってきた。さらに放送局からインターネット経由で大量の番組を配信するサービスも計画されている。このような大量の番組の内容を容易に把握するためのインターフェースが求められるようになり、我々は、テレビ番組を要約した番組紹介を自動生成する研究を進めている。

番組紹介を生成する研究として、河合らは電子番組表(EPG)と番組のクローズドキャプションを利用し、EPGの番組紹介文に類似するクローズドキャプションを抽出し、そのクローズドキャプションに対応する番組映像をつなぎ合わせることで番組映像とそのコメントを生成する手法を提案している[1]。しかし、この手法では、EPGとクローズドキャプションを比較する際、単語の有無を特徴として利用するため、文全体の特徴が抜き出せないことが問題となっている。

本稿では、番組の予告を目的とした短い文章の構文的特徴を学習する。それにより、番組紹介映像の説明文(番組紹介コメント)として利用できる文を番組のクローズドキャプションから Adaboost の機械学習を利用し自動抽出する手法を提案する。NHKで放送された「地球!ふしぎ大自然」を対象とした抽出処理と実験結果について報告する。

2. 番組スポットの特徴

放送される番組には、番組スポットと呼ばれる番組の予告を目的とする短時間に編集された要約映像が存在するものがある。本稿ではこの番組スポットのナレーションを基として、クローズドキャプションからスポット文に相応しい文を抽出することにより、番組紹介コメントを生成する。

番組スポットのナレーションの例を図1に示す。番組スポットのナレーションの各文は番組全体の紹介(以降「紹介」とする)、番組に登場する主人公の特徴(以降「特徴」とする)、興味を惹くために疑問を投げかける謎かけ(以降「なぞ」とする)、番組全体の概要(以降「概要」とする)に分類できる。それぞれの特徴を以下に記す。

- ①「紹介」・・・番組全体の内容が一文で分かるような表現。番組タイトル、主人公や舞台の紹介などが行われる。特徴的な表現が多い。
- ②「特徴」・・・主人公の性格、特徴。舞台となった場所の説明などが行われる。主人公などによって特徴が変わるので、様々な表現が用いられる。
- ③「なぞ」・・・「紹介」や「特徴」を踏まえたうえで、疑問を投げかけている。疑問詞、クエスチョンマークな

どの表現が多い。

- ④「概要」・・・全体のまとめ。番組スポットでは最後にある。全体を捉えるので様々な表現が用いられる。

ふしぎ大自然。次回は日本最大の水生昆虫タガメです。	①
田んぼの中ですばやく動くカエルを意図も簡単に一撃でしとめる姿はまさに水中のハンター。	②
夜になると驚いたことに一気に空へと舞い上がります。	②
しかも何キロも飛ぶことがわかってきました。	②
タガメは一体どこへ飛んでいくのでしょうか?	③
驚くべきタガメの姿を追います。	④

図1. 番組スポットのナレーションの例

3. 番組紹介表現抽出の手順

提案手法では、前節で分類した番組スポットの4項目に対して、その表現が類似するクローズドキャプションをそれぞれ抽出する。人手により番組スポットのナレーションを書き起こし、一文ごとにどの分類項目に該当するかをタグ付けする作業を行った。この分類タグ付き番組スポットナレーションを学習データとし、Adaboost アルゴリズムによって各項目の特徴を学習し、最終的には、生成された最終仮説を新たな番組のクローズドキャプションに適用することにより、番組コメントを抽出する。以下に弱学習器生成と Adaboost による学習処理を示す。

3.1 弱学習器生成

入力テキストを一文ごと構文解析器「南瓜」にかけ、係り受け解析結果を得る。その中から任意の部分木構造を抜き出す。スポット「次回は日本最大の水生昆虫タガメです。」を構文解析した依存構造木と、その依存構造木から抜き出した部分木の一部を図2に示す。依存構造木の各ノードは一つの形態素とし、文節内では先に出現した形態素が後ろの形態素を修飾しているとした構造としている。

この部分木の有無により弱学習機を生成する。例えば、「タガメ」という1単語からなる部分木に対して、弱学習器は、この単語が入力テキストに含まれている場合に+1、含まれていない場合-1を返す関数となる。この処理において、利用するノード数が多いと計算量が膨大になってしまう。しかし、本実験で対象とするテキストデータはスポットのため長文は少ないと考えられる。そこで、選択対象のノード数が少なく、ノード数を制限する。今回はノード数を最大で5個とし、生成する部分木の数を計算可能なものにした。

†中央大学大学院理工学研究科

††NHK 放送技術研究所

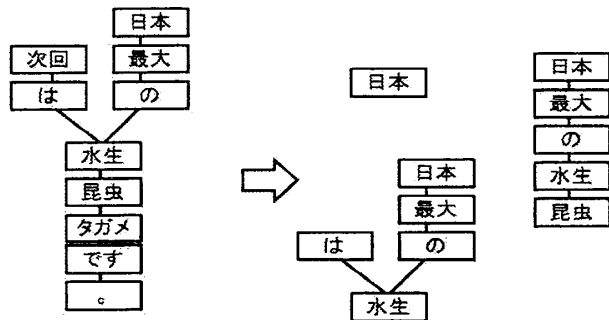


図2. 部分木の抽出

3.2 Adaboost による学習

学習データである番組クロードキャプションから抽出した部分木によって、大量の弱学習器が生成される。この弱学習器は Adaboost の機械学習に利用する。

Adaboost の機械学習器は、集団学習を行う boosting 手法の一種で、分類精度がそれほど高くない弱学習器を多段に組み合わせることにより、高性能の分類器を構成する。入力として与えられる大量の弱学習器に対して、それぞれの程度のカテゴリ能力を持つか判定し、各弱学習器の信頼度を計算する。弱学習器の判定結果 (1 または -1) と信頼度の積を足し合わせ、その値の符号により判定を行う。例えば、「紹介」に相応しい文を抽出する学習器を生成する場合、手作業にて「紹介」に分類された番組スポット文を正例、それ以外を負例として学習する。生成された学習器をクロードキャプションに対して適用し、弱学習器の判定結果と信頼度の積を足し合わせた値の降順に「紹介」の項目に相応しいと判断して抽出する。

4. 実験

NHK で放送された自然番組「地球!ふしぎ大 自然」45 回分を処理対象とした、番組紹介コメント抽出実験を行った。各項目における抽出する数は、1 文登場に対して「紹介」が 3 個、「特徴」6 個、「なぞ」3 個、「概要」3 個とした。この数は各項目の人手によって抽出した正解データの平均個数である。対象とする番組スポットナレーション、クロードキャプションを 6 分割し、クロスバリデーションにより実験を行った。図 1 で示した番組クロードキャプションを対象とした抽出結果の中から最も各項目に近いと判定された文を図 3 に示す。

「ふしぎ大 自然」今日は田んぼで暮らすタガメの謎に迫ります。	① 紹介
タガメには一晩に少なくとも4キロの距離を飛ぶ力が秘められていることがこの調査で初めて分かったのです。	② 特徴
タガメは一体どこに何のために飛んでいくのでしょうか?	③ なぞ
浅くて流れが静かな田んぼは水温が高くメダカの好きなミジンコなどが豊富なのです。	④ 概要

図3. 番組スポットナレーションの例

人手によってクロードキャプション中から番組スポットの各項目に相応しい文を抽出し、この結果を正解データとした抽出結果の評価を行った。また、単語レベルでの抽出評価結果も行った。結果を以下に示す。

表1 部分木レベルでの抽出評価結果

部分木	紹介	特徴	なぞ	概要
適合率	17/195 (8.7%)	42/548 (7.6%)	15/74 (20.2%)	6/192 (3.1%)
再現率	17/156 (10.9%)	42/588 (7.1%)	15/66 (22.7%)	6/135 (4.4%)

表2 単語レベルでの抽出評価結果

単語	紹介	特徴	なぞ	概要
適合率	1/195 (0.5%)	48/548 (8.8%)	1/74 (1.4%)	7/192 (3.7%)
再現率	1/156 (0.6%)	48/588 (8.2%)	1/66 (1.5%)	7/135 (5.2%)

表1、表2より、本手法と単語のみでは「特徴」、「概要」においてあまり差が出なかったが、「紹介」、「なぞ」では、それぞれ大きく本手法において効果が出たと言える。

これは「紹介」、「なぞ」では文全体である程度決まった表現が多く、単語のみでは学習できない特徴的な表現が反映した結果だと考えられる。また、「特徴」、「概要」においてほとんど差がでなかった理由として、多くの種類の表現が多く、文章構造も一定でない。そのため、単語レベル、本手法とで大きな差が出なかったのだと考えられる。

全体的に再現率や適合率が低かった理由としては主人公を特定できなかったことがあげられる。主人公に対して紹介する文章は述べているが、抽出したものの多くに主人公が含まれていなく、主人公とは関係のないことを述べている文章を抽出してしまった。今後主人公を考慮に入れながら抽出できれば再現率や適合率は上がると考えられる。

5. まとめ

本稿では、クロードキャプションから文全体の構文的特徴を用いて、番組を紹介する文章を抽出する手法を提案した。実験により、従来手法に比べて文章の部分木を用いて抽出した方が番組を紹介する文章を抜き出せることを確認したが、依然改善の余地が残されている。今後は、動物、場所、あるいは主人公などを抽象化するなどすることで、再現率、適合率を上げ、さらには抜き出した文章を整形することによって、より番組紹介コメントらしい表現へ変換する予定である。

【参考文献】

- [1] 河合 吉彦,住吉 英樹,八木 伸行:電子番組表を利用した番組紹介映像の自動生成,電子情報通信学会, ITS2006-72, pp.165-170(2007)
- [2] Freund, Y. and Schapire, R.E. A decision theoretic generalization of on-line learning and an application to boosting. Journal of Computer and System Sciences, Vol.55, No.1, pp.119-139(1996)