

# 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.

電気電子情報通信専攻 小林 彰太  
KOBAYASHI Shota

## 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、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。

また、ニュースの原稿を要約生成する研究も進められている。田中らはニュース原稿の構造に着目し全体の概要を説明しているリード文を骨格に用い、残りの文章の本記の表現からリード文と本記の共通単語間を置換、挿入と補完する手法を提案している[2]。しかし、「リード文」、「本記」の構成をとらない文には適用できないため、編集できる対象が限定されるという問題点がある。

本研究では、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生成処理の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。本研究ではこの番組スポットのナレーション(番組紹介コメント)を基に、クローズドキャプションから番組紹介コメント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。

番組紹介コメントの例を図1に示す。番組紹介コメントの各文は主に起承転結で構成されており、番組全体の紹介(以降「紹介」とする)、番組に登場する主題の特徴(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)、番組全体の締め(以降「締め」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、主題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」・・・主題の性格、特徴。舞台となった場所の説明などが行われる。主題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、クエスションマークなどの表現が多い。
- ④「締め」・・・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文から構成される。

ふしぎ大 自然。今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. 番組紹介コメントの例

## 3. 提案手法の概要

本研究では二つの手法を提案する。大きく「抽出」「生成」に分かれる。「生成」と「抽出」はまったく別の処理であり、どちらかの処理結果を利用したり、一方の処理結果により、もう一方の処理が変わることはない。また、本研究の「生成」の実験では4つあるカテゴリーの中でも「締め」文に関して生成を行う。その理由は、「締め」文の特徴にある。「締め」文の特徴は、定型的な表現が多く、かつ、いくつかの内容から要約的に構成されていることが多いため抽出では限界があると考えたためである。

最終的に、「締め」では「抽出」と「生成」でよかったもの選択し評価する（図 2.）.

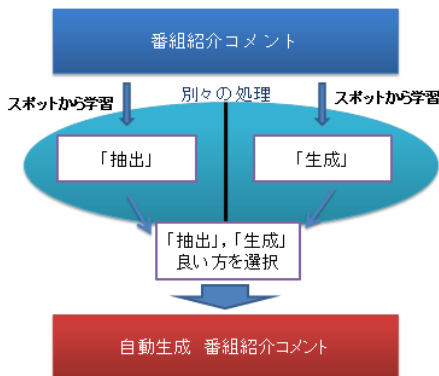


図 2. 提案手法の概要

### 3.1 番組紹介コメント抽出の手順

提案手法では、前節で分類した番組スポットの4つのカテゴリに対して、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（番組紹介コメント）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紹介コメントを学習データとし、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。

#### 3.1.1 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。クローズドキャプション「今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 3 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、「日本」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クローズドキャプション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。今回はノードの数を最大で 5 個とし、生成する部分木の数を計算可能なものにした。

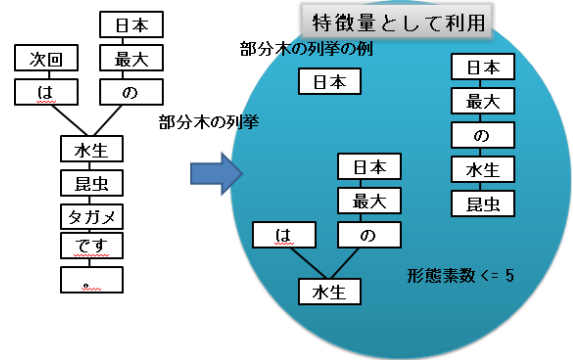


図 3. 部分木の抽出

#### 3.1.2 Adaboost による学習

学習データである番組クローズ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、どの程度の分類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ズ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。

### 3.2 番組紹介コメント生成の手順

今回、「生成」処理では、前節で分類したカテゴリの中で「締め」に対して行う。

「生成」では重要語の抽出を行い。重要語を多く含む文をクローズドキャプションから複数文抽出する。その抽出した文に対して形態素解析を行い、TF-IDF の最も高い値の文を一つ選ぶ（以降、この文をリード文と呼ぶ）。その他の重要語を含む文から重要語にかかる修飾句を抽出し、リード文の重要語と修飾句の被修飾語（重要語）が共通している場合に修飾句を補完する。この処理で番組紹介コメントを生成する。処理の流れを図 4 に示す。

重要語として、今回、「主題」と「頻出品詞」を採用する。「主題」は TF-IDF 法により番組のクローズドキャプションから抽出。「頻出品詞」はカテゴリ別の各品詞の出現回数 がクローズドキャプション

より出現する割合が高い品詞を抽出する。

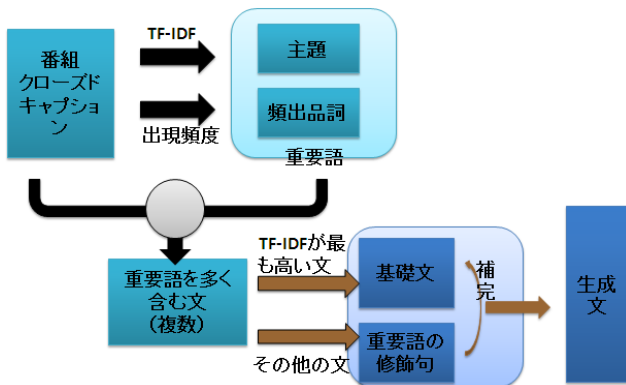


図4 生成処理の流れ

### 3.2.1 TF-IDF の算出

TF-IDF は TF と IDF という二つの指標を掛け合わせて計算される。TF とは出現頻度のことであり、多く出現する単語を大きくする。IDF は逆文書頻度のことであり、多くの文書に出現する単語を小さくする重み付けである。以下の式で計算される。

$$tf(w) = n(w) \quad (1)$$

$$idf(w) = \log\left(\frac{N}{df(w)}\right) \quad (2)$$

ここで、 $n(w)$  は単語の出現頻度を表し、 $|N|$  は全文書数、 $df(w)$  は単語  $w$  が出現した文書数を表している。

式(1)、式(2)を掛け合わせることで、出現頻度と逆文書頻度とを考慮に入れた重み付けが可能となる (TF-IDF, 式(3))。

$$tfidf = tf(w) \cdot idf(w) \quad (3)$$

## 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と生成実験を行った。また、その抽出、生成した結果 5 番組を被験者 12 人に対して、主観評価実験を行った。

### 4.1 抽出生成実験

抽出実験の各カテゴリーにおける抽出する文数は、「紹介」が 1 文、「特徴」3 文、「なぜ」1 文、「締め」1 文とした。この数は各カテゴリー別の番組紹介コメントの平均文数である。対象とする番組紹介コメント、クローズドキャプションを 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ズドキャプションを対象とした抽出、生成実験の結果の中から各項目に近いと判定された文を図 5 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。	紹介
タガメには一晩に少なくとも 4 キロも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。	特徴①
父親となった オスには大切な仕事 待っているのです。	特徴②
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。	特徴③
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？	なぜ
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。	締め

図 5. 番組紹介コメント抽出の例

生成は「締め」のみに対して行い、1 文生成した。「生成」実験の結果の例を図 6 に示す。

田んぼで暮らすタガメの持ちながら一晩に何キロも飛ぶ 意外な秘密に迫ります。	締め (「生成」)
--	--------------

図 6. 番組紹介コメント生成の例

正解データとした抽出結果の評価を行った。また、既存手法の河合らの手法による比較実験も行った。結果を以下に示す。

表 1 「抽出+生成」実験の結果 (「本手法」)

	BLEU	ROUGE	F値
紹介	0.18	0.13	0.15
特徴	0.18	0.11	0.14
なぜ	0.33	0.28	0.30
締め (抽出)	0.12	0.10	0.11
締め (生成)	0.27	0.26	0.27

表 2. 「単語」の結果 (「単語」)

	BLEU	ROUGE	F値
紹介	0.12	0.08	0.10
特徴	0.21	0.09	0.12
なぜ	0.14	0.06	0.09
締め	0.15	0.08	0.11

表 1、表 2 より、「本手法」と「単語」のみでは抽出手法における「締め」と「特徴」ではほとんど差が出なかったが、その他のカテゴリーでは「本手法」が上回った。

これは「紹介」、「なぜ」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」

や抽出手法における「締め」において、ほとんど差がでなかった理由として、文章構造は一定だが、多くの種類の表現が多い。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。

また、「締め」において「生成手法」が「単語」や「抽出手法」より良好になった理由も「締め」の文章構造は一定だが、多くの種類の表現が多いという特徴から良好な結果になったと考えられる。

全体的に F 値が低かった理由としては抽出手法において主題とは関係のないことを述べている文章を多く抽出してしまった。今後主題を考慮に入れながら抽出できれば F 値は上がると考えられる。

## 4.2 主観評価実験

抽出と生成した文章（ランダムに取り出した 5 回放送分）で被験者 12 人の学生に対して、意味が十分わかるものと、わからないもの（評価）を 5 段階（とてもわかりやすい 5 点～とてもわかりにくい 1 点）で評価してもらった。

また、どのような内容か簡単に記述してもらった。記述してもらったものと、実際に放送された番組紹介コメントから人手で抽出した主題と内容を比較し、主題が一致していたら 1 ポイント、主題かつ内容が一致していたら 2 ポイントとした。

また、「締め」は結果が良好だった「生成」の結果を採用した。この手順で点数を算出した結果を表.3、表 4.に示す。

表.3 読みやすさの評価

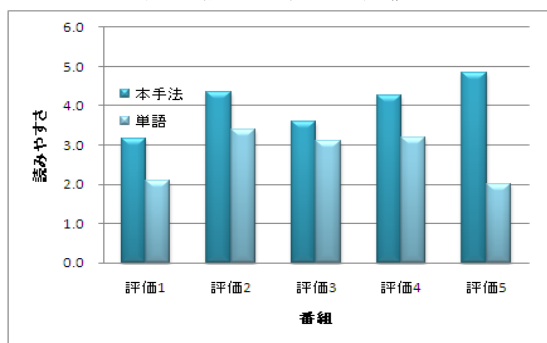
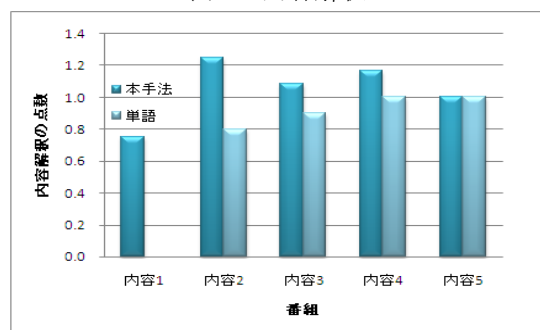


表.4 内容解釈



「本手法」と「単語」の二つを比較すると、《理解の容易さ》で「本手法」のほうが平均で 1.28 上回り、《内容理解》で平均 0.26 上回った。また、「本手法」は 91.2[%]の《内容理解》の項目で主題を抜き出したのに対し、『単語』は 70.0[%]だった。また、被験者からは『「本手法」の方が文章の流れがいい。』、『「本手法」の方が、何が言いたいか、何に対して言っているかはっきりしている』『「本手法」は重複文が『単語』に比べ少ない。』などの意見が出た。このことから主観評価において「本手法」の方が良好だったといえる。『単語』のみ学習するより、『部分木』で学習したほうが、文の特徴を考慮に入れることができるため重複が少なく、文の流れが考慮に入れることができたためこのような結果になったのだと考えられる。

## 5. まとめ

本研究では、クローズ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、生成する手法を提案した。また、主題や頻出品詞を中心とした生成手法も提案した。また、カテゴリ別に学習することで文章を作り出すことに成功した。

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。また、生成手法も良好な結果が出ることを確認した。しかし、今後改善の余地が残されている。今後は、動物、場所、あるいは主題などを抽象化するなどすることで、F 値を上げ、さらには整形、生成することによって、より番組紹介コメントらしい表現を自動生成するよう取り組むことが必要だと考えられる。

今後、この技術が新たな発達をし、本研究であげた課題を解決することで番組紹介コメントに限らず多くの利用価値のある文章を自動生成することに期待される。様々な文章の自動生成ができるようになり、多様な番組や動画のかかわり方を助ける技術となることを期待したい。

## 【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] 田中 英樹,木下 明德,後藤 功雄,熊野 正,加藤 直人:放送ニュース要約のための表現置換,言語処理学会, NLP2007