

番組のシーン集合へのラベリングの検討

三浦 菊佳[†] 松井 淳[†] 山田 一郎[†] 後藤 淳[†] 宮崎 太郎[†] 宮崎 勝[†] 住吉 英樹[†]NHK放送技術研究所[†]

1 はじめに

近年、動画コンテンツを扱うインターネット・サービスが急増しており、大量のコンテンツをわかりやすく提示する必要性が増してきている。放送局においても、過去に番組で扱ったテーマの分析や、NHK オンデマンド[1]などインターネットを経由した視聴者向け動画配信サービスのために、これまで放送した番組を整理して一覧することが求められている。また、特定シーンのみをピンポイントで見たいといったユーザの要望も高まっている。

過去の放送番組を内容に基づいて構造化し検索する手段として、我々はテレビ番組マップを提案している[2]。これは、過去の放送番組に対して、単語を介してリンク付けしたもので、例えば、「きょうの健康 → 疲れ目 → 予防法 → ブルーベリー → きょうの料理」のように、ある番組（きょうの健康）から別の番組（きょうの料理）を辿ることができる。このテレビ番組マップは、検索や推薦を目的とした内部処理用として構築され、可視化用途ではないため、全ての番組を俯瞰することが困難であった。また、処理対象が番組単位で内容を記述した EPG（電子番組表）の番組概要文であるため、番組内の一部分やある特定のシーンに対応することができなかった。

そこで、番組間の関係の全体像から細部までを可視化することを目的としたテレビ番組マップを提案する。意味的なまとまりで番組をクラスタに分け、そのクラスタ内容を表すラベルを付与する。本稿では、ラベルを自動的に付与する手法について述べる。



図1. テレビ番組マップのインターフェース
(左: 全体図、中央: ズームアップ画面、
右: 円周にシーン、内側にクラスタ)

A Study on Labeling to the Scene Sets of TV Programs

Kikuka MIURA[†] Atsushi MATSUI[†] Ichiro YAMADA[†] Jun GOTO[†] Taro MIYAZAKI[†] Masaru MIYAZAKI[†] and Hideki SUMIYOSHI[†][†] Science and Technology Research Laboratories,
Japan Broadcasting Corporation

2 テレビ番組マップとクラスタのラベル

新しい番組マップのインターフェース画面の例を図1に示す。図1左に示したテレビ番組マップの出力画面は、シーンの集合をクラスタリングした結果を可視化したものである。ユーザは、まず、過去の放送番組に出現した話題の全体像をこのインターフェースで俯瞰し、次に、個々のクラスタに付与されたラベルを手がかりにして興味を持つクラスタを選択する（図1中央）。最終的に、興味を持つ話題で絞りこまれた番組シーンの集合を提示する（図1右）。ここで各クラスタに付与したラベルがクラスタを選択する手掛かりとなる。クラスタのラベルは複数の語で構成され、内容を端的に表すことに加え、他のクラスタとの意味的な差異が明示的であることが求められ、クラスタの取捨選択の過程において後者は特に有効と考えられる。

黒田ら[3]は、タイトルボタンを用いてクラスタにタイトルを生成する手法を提案しているが、単語の頻度情報を用いて重要語を選出しているため、意味的な類似性を考慮していない。また、奥村ら[4]は、ニュース記事クラスタへのラベル生成手法を提案しているが、上位概念を獲得する際に、別途、オントロジーを必要とする。

提案手法では、オントロジーなどの外部の知識を必要としない手法として、単位時間で分割したクローズドキャプションの集合に対して、多段階でクラスタリングを適用し、得られた各々のクラスタにラベルを付与する。この際、単語の分散表現によりクラスタを代表する語を抽出し、さらに、他のクラスタとの差分を表す特徴を利用する。実験では、NHK 総合テレビで放送した1年分の番組を対象とし、提案するラベル付与手法の有効性を確認した。

3 提案手法

番組マップ生成は、番組のナレーションや出演者の発話内容が文字化されたクローズドキャプションを入力として、以下のStep1～Step5の手順で行う。

Step1: セグメント生成

クローズドキャプションの集合を固定長の時間単位で分割してセグメントを生成する。

Step2: 各セグメントのベクトル作成

各セグメントの特徴を表すベクトルを作成する。この処理では、クローズドキャプションに出現する単語をword2vec[5]により200次元の単語ベクトルで表現し、各セグメントに含まれる単語ベクトルの和をセグメントベクトルとする。

Step3: 多段クラスタリング

まず、全セグメントを k -means 法により k 個のクラスタに分類する。この際、クラスタ間の類似度は Step2 で

生成したセグメントベクトルのコサイン類似度により計算する。次に各クラスタを再度 k 個に分類し、この処理を繰り返す。この結果、Step1 で生成した全セグメントが多段のクラスタリング処理により分類される。

Step4: クラスタに対するラベリング

Step2 で生成したセグメントベクトルを利用する。まず、クラスタに属するセグメントベクトル集合の平均を計算してクラスタベクトルを生成する。このクラスタベクトルと Step2 で算出した単語ベクトルとのコサイン類似度を計算し、類似する単語ベクトル n 個を、クラスタを代表する代表語として抽出する。この際、代表語の優先度はコサイン類似度の降順とする。次に、他のクラスタとの違いを明確化する差分語を抽出する。1つのクラスタに注目すると、Step3 の k -means 法によるクラスタリング処理により、各クラスタには $(k-1)$ 個の兄弟クラスタが存在する。そこで、注目するクラスタベクトルと、一つの兄弟クラスタベクトルとの差分を取り、その差分ベクトルに類似する単語ベクトルを m 個抽出する。 $(k-1)$ 個の兄弟クラスタに対してこの処理を行い、複数のクラスタにおいて重複して出現する単語を差分語とする。最後に、クラスタの代表語と差分語を用いてクラスタラベルを決定する。代表語と差分語が共通する場合、代表語の優先度の降順にクラスタラベルとする。ラベル数が N に達していない場合は、残りのラベルを代表語から優先度の降順に選択する。

Step5: クラスタベクトルの次元圧縮

クラスタ間の距離の比を保持しつつ、クラスタベクトルを 200 次元から 2 次元に圧縮する。この処理により、図 1 に示す可視化を実現する。

4 実験および考察

4-1 セッティング

ラベル獲得手法の有効性を検証するため、NHK で放送された番組（2013 年総合テレビ 1 年分、計 11,378 番組）を対象に実験した。形態素解析には MeCab[6]を使用し、セグメント数は、番組を 1 分ごとにシーン分割した結果、312,480 となった。 k -means におけるクラスタの分割数 $k=5$ とし、クラスタリングを 5 段階行った。ラベル獲得の際のパラメータとして、代表語数 $n=30$ 、差分語数 $m=5$ とした。また、セグメントベクトル作成に際して処理対象とする単語は名詞に限定し、代名詞や記号などのストップワードは除外した。

評価では、著者ではない 1 人のアノテータが、最下位層の 3,125 個のクラスタから 250 個を選択し、出力したラベルに対して、そのクラスタを表すラベルとして相応しいか否かの 2 値を与えた。ベースライン手法として、tf-idf によりクラスタ内の単語に重みを与え、その降順にクラスタラベルとして抽出した。

4-2 実験結果と考察

評価結果を図 2 に示す。ラベル数 N が 1 から 10 の間における、代表語と差分語を利用した提案手法、比較手法として、代表語のみを利用した手法、ベースライン手法における各正解率を表す。ラベル数が小さい設定 ($N=1\sim4$) においては、概ねベースライン手法が提案

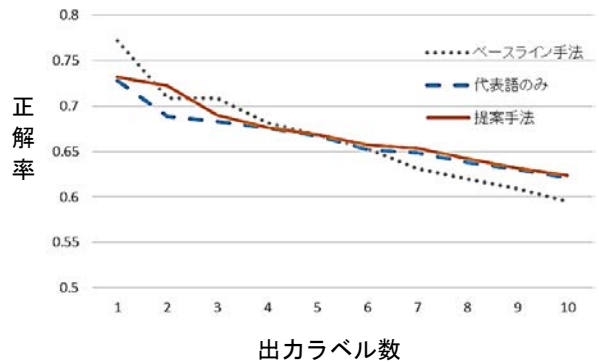


図 2. 評価結果

手法を上回る正解率であったが、ラベル数が 5 以上では提案手法がベースライン手法を上回った。今回の可視化システムでは、各クラスタのラベル数 $N=10$ 程度を想定しており、提案手法は有効であるといえる。また、すべての出力ラベル数 N において、代表語のみから獲得する手法よりも、差分語を考慮した手法の方が、アノテータが相応しいと判断する結果となった。

提案手法では、クラスタに属するテキスト中に出現しない単語もラベルとして獲得できる。実際に、「価格競争」「市場」「成長」の記述があるクラスタにおいて、「コストダウン」「ビジネスチャンス」が、「皇居」「パレード」「モーニング」の記述があるクラスタにおいて「セレモニー」が獲得された。

提案手法で不正解と判定された例としては、歌番組やドラマなどが含まれるクラスタで、歌のタイトル、人名、役名がラベルとして取り出され、そのラベルがクラスタ中に存在しないケースが多く見られた。人名などは同じような傾向の単語ベクトルとなりやすいため、誤って出演していない人名が選択されてしまった。今後、クラスタやラベルが属するドメインによって処理を変えていく必要があると考えられる。

5 まとめ

本稿では、放送された大量の番組の各シーンを多段階でクラスタリングし、それぞれのクラスタに対してラベルを付与する手法を提案した。字幕放送の番組 1 年分 11,378 番組のクローズドキャプションを対象として獲得したラベルを評価し、ラベル数 $N=5\sim10$ で、クラスタの代表語と差分語を利用する提案手法がベースライン手法と比較して有効であることを示した。今後、番組分析システムや番組推薦システムなどのインターフェースに応用していくことを検討している。

参考文献

- [1]NHK オンデマンド. <https://www.nhk-ondemand.jp/>
- [2]三浦ほか. 単語間の意味的關係を用いた番組リンク生成. 信学技報 NLC2014-42
- [3]黒田ほか. 文書クラスタリングにおけるクラスタタイトルの自動生成. FIT2008 E-060
- [4]奥村ほか. ニュース記事クラスタの知的ラベル付け. DEIMForum2015 D1-2
- [5]T.Mikolov, et al. Distributed Representations of Words and Phrases and their Compositionality. Proc. NIPS2013
- [6]MeCab. <http://taku910.github.io/mecab/>