

# 内容的特性に基づくTVCMのアピール特徴抽出

澤田 悠治<sup>1)</sup> 北山 大輔<sup>2)</sup> 角谷 和俊<sup>1)</sup>

1) 関西学院大学 2) 工学院大学

## 研究背景 TVCMでアピールされている特徴をユーザは評価しているのか?

### ■ TVCM (2013年1月放送)



「急須への挑戦」 「このお茶は、急須で入れた緑茶の味?」 「急須で入れた緑茶と間違えました。」 「にごりは、旨み。」

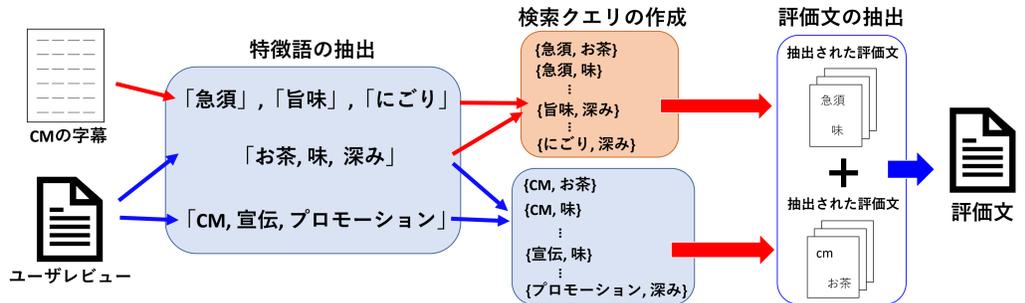
### ■ ユーザレビュー

ユーザA 「やっぱり本当にお茶とは味が全然違う」

ユーザB 「急須で入れたように濃いとの触れ込みですが、全然濃くない」

... ユーザZ 「にごりの旨みと書いてあるが、あまり旨みは感じない。」

## アプローチ TVCMの字幕からCMから商品を購入した投稿を抽出



- TVCMの頻出単語及びユーザーレビューから属性語を抽出
- 抽出したキーワードを用いて検索クエリを生成
- 検索クエリにより商品に関する評価文を抽出

## 提案手法 Step.1 特徴語の抽出

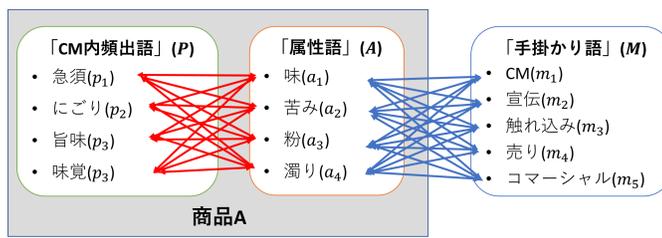
- CM内頻出語
  - ✓ TVCMの字幕データで全期間、各CMで頻出する副詞、一般名詞全期間
$$Entired\ Frequency_i = \frac{cf_i}{N}$$

該当期間

$$Limited\ Frequency_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}}$$
  - $cf_i$ : 単語*i*が出現するCM数
  - $N$ : 対象商品におけるCMの総数
  - $n_{i,j}$ : CM*i*で単語*j*が出現する頻度
  - $\sum_k n_{k,j}$ : CM*i*における総単語数
- 属性語
  - ✓ 対象商品の全投稿に含まれる一般名詞
$$Item\_Frequency_i = \frac{n_i}{\max(n_{k \in D})}$$
  - $n_i$ : 単語*i*の出現頻度
  - $\max(n_{k \in D})$ : 最大出現頻度
- 手掛かり語
  - ✓ "CM", "広告", "宣伝", "触れ込み"と類似する名詞
  - ✓ Word2Vecで全投稿の学習データからコサイン類似度を元に抽出

## Step.2 検索クエリの作成

3つの特徴語の全ての組み合わせから検索クエリを作成



- CM内頻出語と属性語による検索クエリ
  - CM内頻出語(P), と属性語(A)の直積集合:  $P \times A$
  - 該当期間は放送後3ヵ月間を指定
- 手掛かり語と属性語による検索クエリ
  - CM内頻出語(M), と属性語(A)の直積集合:  $M \times A$

## Step.3 評価文の抽出

作成した検索クエリを元に評価文を抽出

- CM内頻出語, 属性語
 
$$Query(p_1, a_1) = U_{i \in D} \{d_i | p_1 \in d_i, a_1 \in d_i\}$$

$$CM\_Set = U_{i \in P, j \in A} Query(p_i, a_j)$$
- 手掛かり語, 属性語
 
$$Query(m_1, a_1) = U_{i \in D} \{d_i | m_1 \in d_i, a_1 \in d_i\}$$

$$Media\_Set = U_{i \in M, j \in A} Query(m_i, a_j)$$
- $U_{i \in D} \{d_i | p_1 \in d_i, a_1 \in d_i\}$ : CM内頻出語( $p_1$ ), 属性語( $a_1$ )を含む投稿
- $U_{i \in D} \{d_i | m_1 \in d_i, a_1 \in d_i\}$ : 手掛かり語( $p_1$ ), 属性語( $a_1$ )を含む投稿
- 評価文の抽出
 
$$Review = CM\_Set \cup Media\_Set$$

## 実験

- 対象データ
  - 不満データで飲料品Aに関する投稿
  - 全投稿: 161件
- CMデータ
  - 2007-2016年で放送されたCMから作成
- CM内頻出語
  - 全期間は0.45, 該当期間は0.2以上を使用
  - 商品名はノイズとして除去
- 手掛かり語
  - Word2Vecで学習した不満データの全投稿データ
- 属性語
  - 出現頻度(最大頻度を1.0)が0.1以上

## 飲料品Aの特徴語

- CM内頻出語
  - 全期間
    - ['旨み', 0.526], ['急須', 0.474], ['にごり', 0.842], ['宇治', 0.526], etc
  - 該当期間
    - [1, '2007/10', 'にごり', 0.29][2, '2007/10', '姿', 0.25][3, '2008/', 'にごり', 0.4]
    - [4, '2009/', 'にごり', 0.43][4, '2009/', 'にごり', 0.36][4, '2009/', '本物', 0.21]
    - [5, '2009/', 'あし', 0.31][5, '2009/', 'あじ', 0.23]
- 手掛かり語
  - CM, コマーシャル, 宣伝, pr, プロモーション, 広告, 触れ込み, ふれこみ, 謳い文句, うたい文句, 言説, キャッチコピー, キャッチフレーズ
- 属性語
  - ['ペットボトル', 50], ['味', 18], ['ボトル', 17], ['ラベル', 14], ['音', 10], ['苦味', 9]
  - ['日本人', 9], ['味覚', 9], ['濁り', 8], ['茶', 8], ['最後', 8], ['商品', 7], ['部分', 7]
  - ['キャップ', 7], ['香り', 6], ['底', 6], ['気持ち', 6], ['色', 6], ['葉', 6], ['段ボール', 6]
  - ['見た目', 5], ['容器', 5], ['粉', 5], ['指', 5], ['箱', 5]

## 飲料品Aの抽出結果

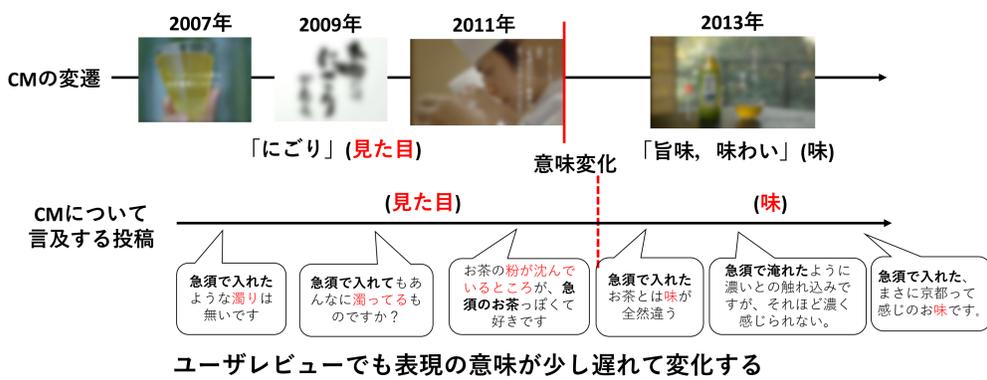
	① 手掛かり語	② CM内頻出語	③ 両方
抽出件数	11	11	17
正解評価文	7	10	12
適合率	0.64	0.91	0.82
再現率	0.41	0.59	0.82
F値	0.50	0.71	0.82

### 表の計算内容

- ✓ 該当期間からのCMから抽出された投稿は無し
- ✓ 全投稿の正解評価文: 17件
- ①: (手掛かり語, 属性語)で抽出した結果
- ②: (CM内頻出語, 属性語)で抽出した結果
  - 該当期間からの抽出なし
- ③: ①, ②の両方の組み合わせから抽出した結果

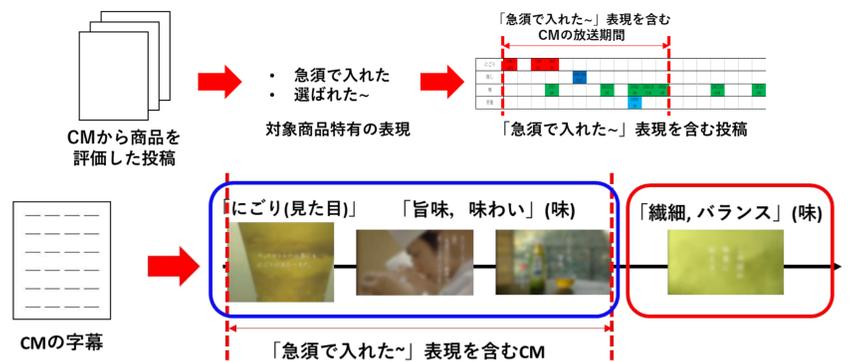
## 内容的特性①: アピール内容の変化

商品特有の表現の意味が変化しているポイントが存在



ユーザーレビューでも表現の意味が少し遅れて変化する

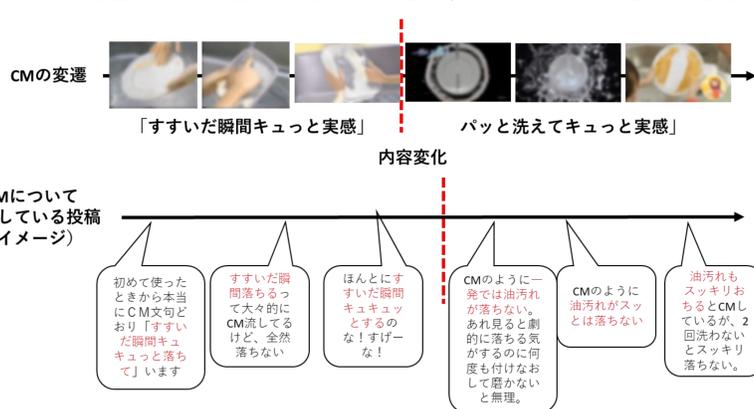
## 特有の表現がされた放送期間でアピールされた特徴の変化を比較



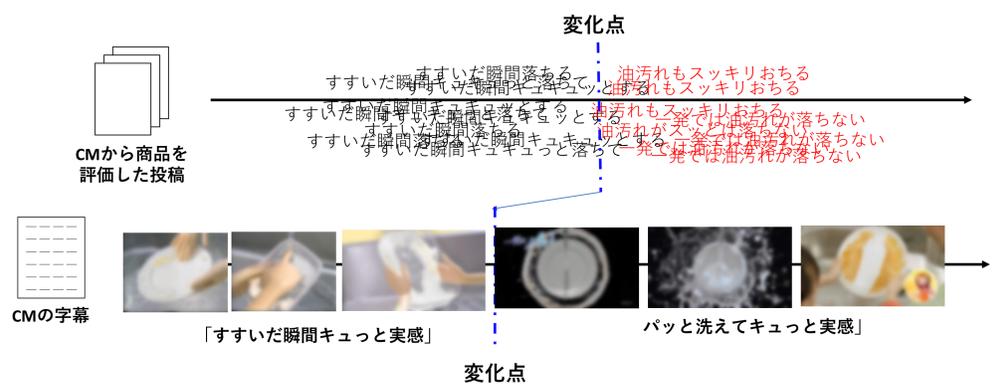
放送期間以降のTVCMで特有の表現がある場合、商品の表現として認知された?

## 内容的特性②: 言及の変化

CMの変遷に応じて言及内容が変わると、レビュー上の言及内容も変化



## TVCMの言及内容の変化とユーザーレビューの言及内容の変化を比較



ユーザーレビューのCMの言及内容から変化点を取得