



テレビCMの視聴覚特徴量がもたらす 消費者の態度変容について

同志社大学 文化情報学部 統計科学研究室 4年

深田大登

目次

1, 研究背景

- ・ コモディティ化の進展と対策
- ・ 感覚マーケティングの関連研究

2, 研究方針

- ・ 研究目的
- ・ 使用データ
- ・ 分析方針

3, 基礎分析①

- ・ 基礎集計
- ・ 広告効果量の推定
 - ・ セミパラメトリックなDID法
 - ・ 傾向スコアモデリング
 - ・ 広告効果量の推定結果

4, 基礎分析②

- ・ 動画データの前処理
- ・ 特徴量抽出

5, 本分析

- ・ 時系列クラスタリング
- ・ 正準相関分析
- ・ 視聴覚変数の前処理
- ・ 分析結果の解釈

6, 提言

- ・ 本研究の展望
- ・ 本研究の課題

参考文献

研究背景

コモディティ化の進展とその対策

研究方針



1

デジタル化・技術発展に伴い**コモディティ化**が進展している。
(※**コモディティ化**:価格競争による、商品の価格低下) [13]
→**価格**や**品質**以外で商品の差別化が必要

2

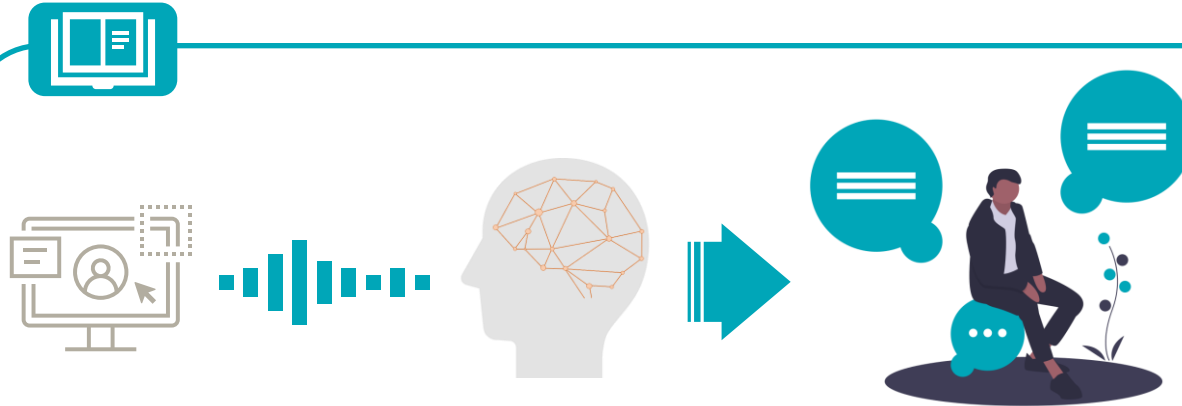
感覚マーケティングでの心理的価値による差別化が有効。 [16]
(※**感覚マーケティング**:消費者の五感を刺激し、その知覚や行動に影響を与えるマーケティング) [3]

特に広告活動では、消費者が受け取る**感覚刺激**と、それに伴う**広告効果**を理解することが重要 [3]

研究背景

感覚マーケティングの先行研究と課題

研究方針



Krishnaら(2016)[4]やTaoら(2020)[7]では、
広告の持つ感覚刺激が消費者に与える
態度変容を広告効果としている。

① Krishnaら(2016) **複数の感覚刺激**の受け取りは考慮されていない。

② Taoら(2020) テレビ広告では**視覚刺激**と**聴覚刺激**が広告効果に最も重要な要因である。
広告効果間の関係は考慮していない。
推定モデルであるため、**解釈性**が低い。
(→分析結果から広告の**改善点**を捉えにくい。)

本研究の目的と期待される成果

本研究の目的

広告効果のあるテレビCMが持つ視覚刺激と聴覚刺激の統合的特徴を探る。

期待される成果

広告効果の期待できる
視聴覚刺激を提案

広告クリエイティブの指標の作成

使用データ

NRI提供データ

① アンケートデータ

2500サンプルのシングルソースデータ

個人属性，チャンネル利用，消費価値観，
商品別の購入実態，
テレビ番組の視聴有無(2021/1/23~4/3)
などが分かる

② 出稿データ

テレビ番組別にどのCMが出稿されたか分かる

CM動画データ

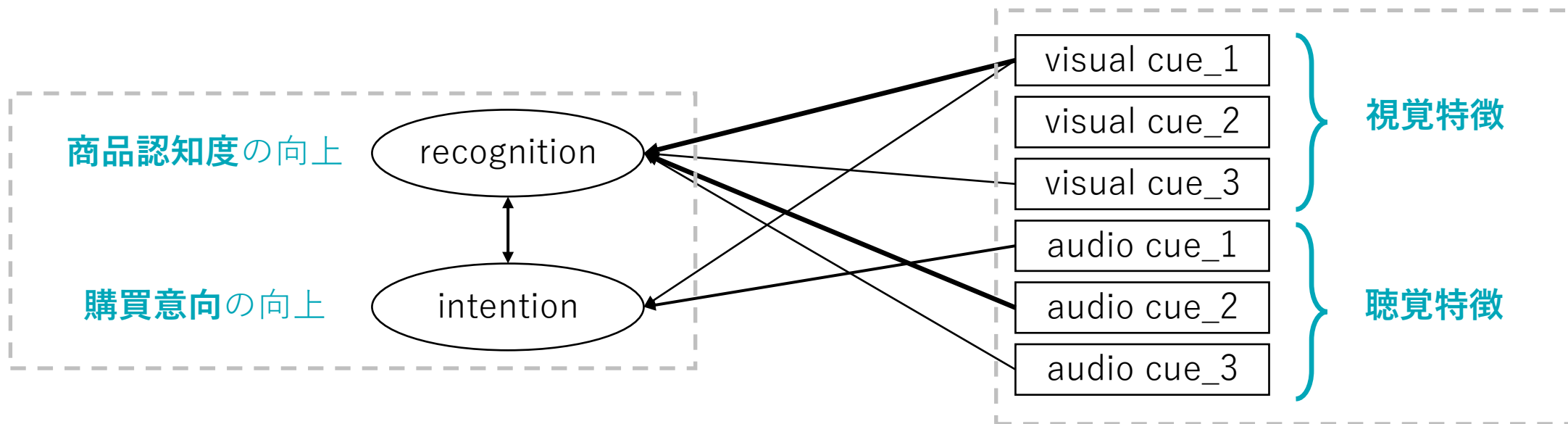
CMの放送時期・商品名・企業名から実際のCM動画を特定し，公式ホームページ・公式YouTubeなどから収集。

(※) 分析に使用する動画は，
アンケートデータの調査期間に出稿されており，
サンプルの10%(250人)以上が視聴しているものとする。（詳細は後述）

分析概要

本分析：CMの広告効果を目的変数，視聴覚特徴量を説明変数とするモデリング

(アウトプットイメージ)



基礎分析①：各CMの広告効果量を推定

基礎分析②：CM動画から視聴覚特徴量の抽出

基礎分析①

基礎分析②

調査対象者の各CM視聴有無を集計

Step①

アンケートデータから商品別の調査期間を調べる。

**Step②**

調査期間に放送されたテレビ番組中に、どのCMが何回出稿されているか調べる。

**Step③**

調査対象者がそれぞれの商品に関するCMを視聴していたかを調べる。

**Step④**

視聴のあるCMのうち**15秒**のものを抽出する。（※CM動画を時系列データとして扱うため）

基礎分析①

広告効果量の推定

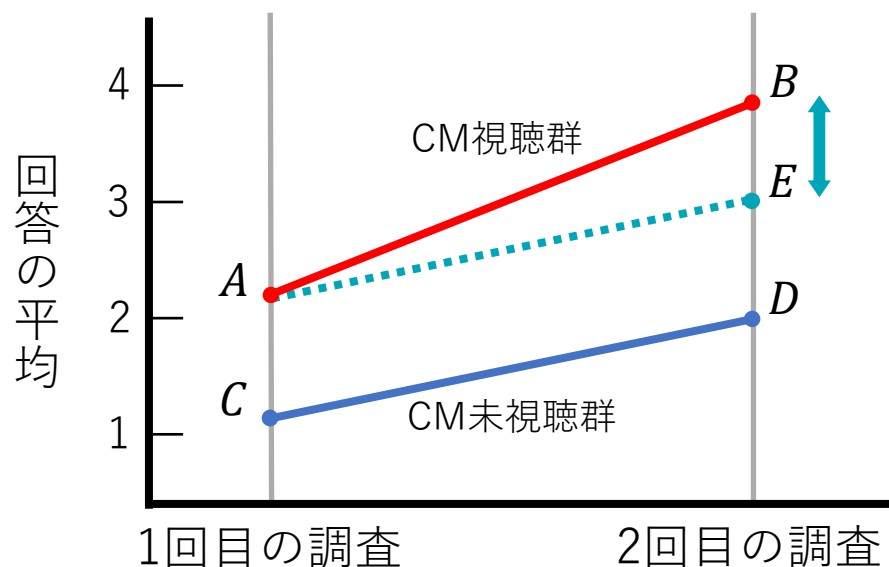
基礎分析②

・アンケートデータの購買意向・商品認知の質問を用いて差分の差法にて広告効果を定義する。

質問は一定の調査期間を設けて、同一のサンプルに対し2回行われている。

→調査期間で質問対象商品に関連するCMを視聴したか否かで回答結果がどのように変化したか調べる。

差分の差法(以下, DID法と呼ぶ) … 2群の1回目と2回目の調査の平均の差を比較する手法



$$\text{DID推定量} = (B - D) - (A - C)$$

ただ, 2つの仮定が前提になっている。

・平均トレンド仮定

CM視聴者がもしCM視聴をしなかったらCM未視聴群の結果と同じになる。

・共通ショック仮定

調査期間において対象者は回答について同じ影響を受けている。

基礎分析①

セミパラメトリックなDID法

基礎分析②

アンケートデータに存在する調査対象者の属性を共変量として、
DID法より仮定が緩い、**セミパラメトリックなDID推定量**(\widehat{DID})[12]を広告効果量とする。

$$\widehat{DID} = \frac{1}{N} \sum_{i=1}^N \frac{Y_{i1} - Y_{i0}}{P_w} \frac{W_i - e(X_i)}{1 - e(X_i)}$$

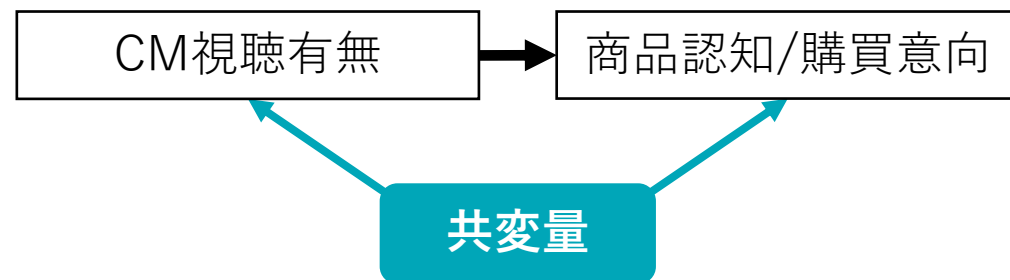
N : 調査対象者数, X_i : 調査対象者 i の共変量

Y_{it} : 調査対象者 i の t 時点の購入意向(商品認知度)

$P_w = \frac{\sum_{i=1}^N W_i}{N}$: 視聴割合, $e(X_i) = \text{Pr}(W_i = 1|X_i)$: 傾向スコア

$t = \begin{cases} 0 : 1\text{回目の調査} \\ 1 : 2\text{回目の調査} \end{cases}$, $W_i = \begin{cases} 0 : \text{サンプル}i\text{のCM視聴あり} \\ 1 : \text{サンプル}i\text{のCM視聴なし} \end{cases}$

共変量とは...



傾向スコアとは...

調査対象者が2群のいずれかに所属する確率のこと。

調査対象者の共変量を、所属確率として1つの値にまとめる。

基礎分析①

基礎分析②

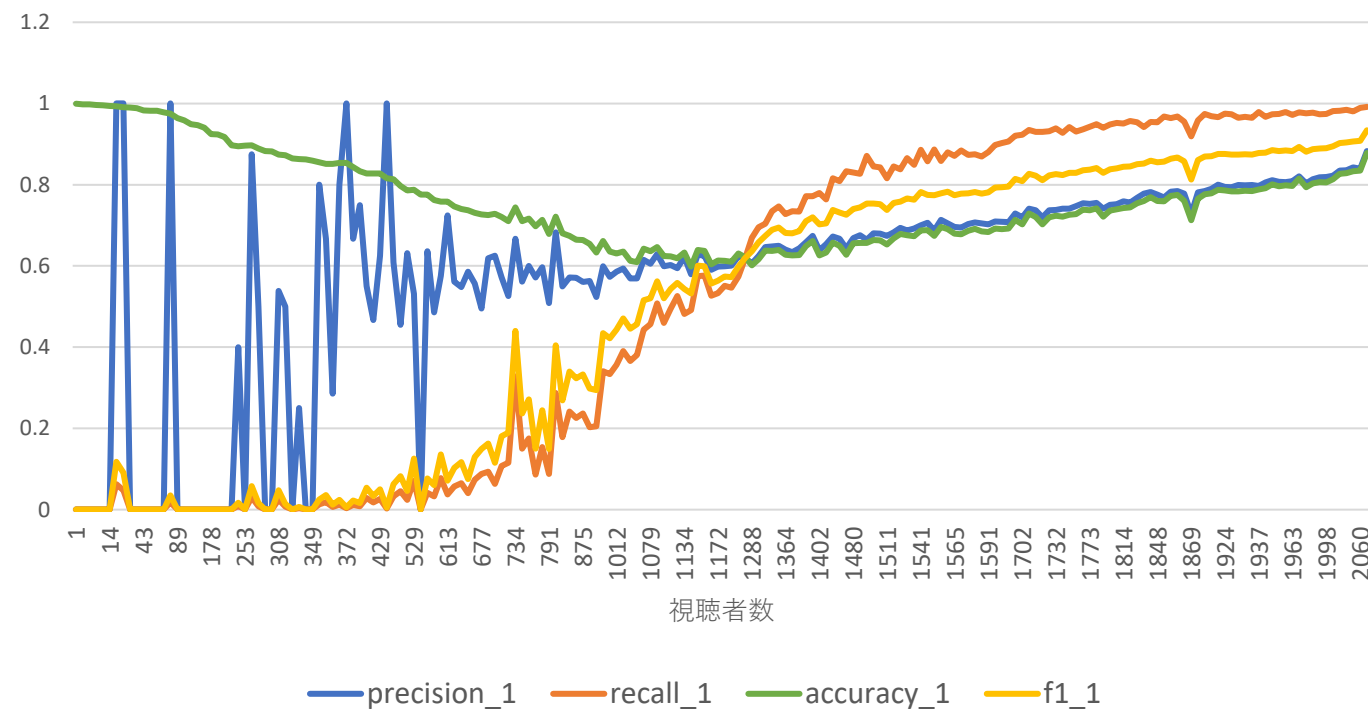
傾向スコアモデリング

推定手法：ロジスティック回帰分析

目的変数：該当CMを視聴したか(0,1)

説明変数(共変量)：性別，年齢，未婚既婚，子供の有無，消費価値観，チャンネル利用頻度．[11]

傾向スコアモデリングの精度比較（視聴者数別）



傾向スコアモデリングでの課題

視聴者が少ないような不均衡データで
適合率の悪さが目立つ。

→CMによって傾向スコアの信頼度が変わる。

→誤差関数やサンプリングでの工夫が必要。

基礎分析①

不均衡データのSMOTEによる対処

SMOTE：不均衡データのデータ数が少ないカテゴリについて、
近傍データを用いた**オーバーサンプリング**

cm_code：1604773の例

不均衡データ) 視聴者：272，非視聴者：2118

元データ	0	1
0	1286	832
1	98	174

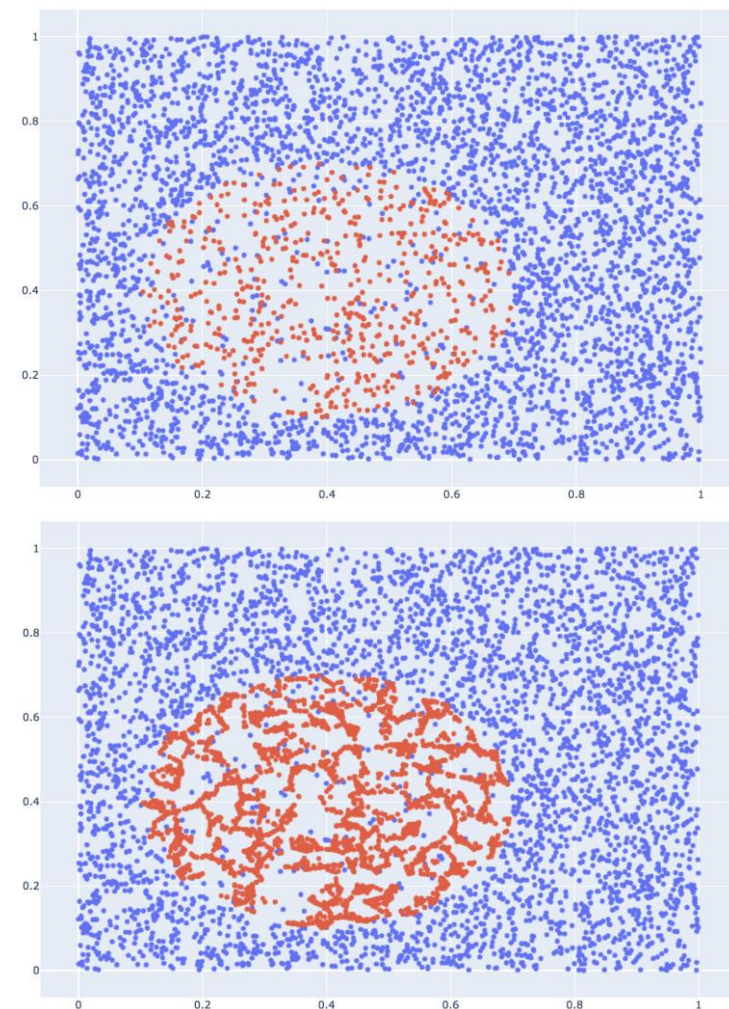
正解率：0.61
適合率：**0.17**
AUC：0.67

SMOTE) 視聴者：2118，非視聴者：2118

実装後	0	1
0	1499	619
1	648	1470

正解率：0.70
適合率：**0.70**
AUC：0.67

不均衡データとSMOTE実装のイメージ



基礎分析①

広告効果量の推定結果



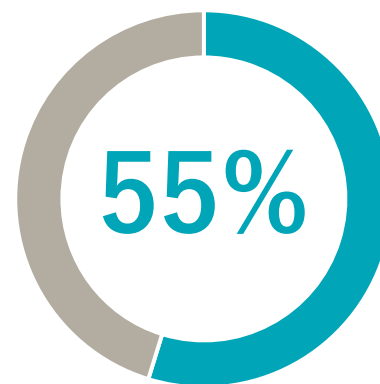
基礎分析②

広告効果 : セミパラメトリックなDID推定量
 対象項目 : 商品認知度, 購買意向
 傾向スコアの推定 : ロジスティック回帰
 共変量 : 性別, 年齢, 未婚, 子供有無, 消費価値観, チャンネル利用頻度
 推定対象 : **視聴者数が全体の10% (250人)以上**のCM
 不均衡データへの対処 : SMOTEによるオーバーサンプリング

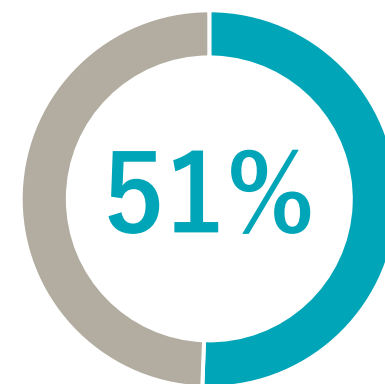
cm_code	watcher	item_name	company	recognitio	intention
1592093	988	アマゾン	Amazon	-0.16863	-0.13811
1604997	1860	アマゾン	Amazon	0.196262	0.191899
1597802	1967	ENEOSで	JXTGエネ	-0.11939	-0.10488
1601178	1762	au PAY	KDDI	-0.13435	-0.08806
1595547	1912	povo	KDDI	-0.12033	-0.11067
1606142	1952	povo	KDDI	-0.25116	-0.18507
1588651	1814	LIFULL H	LIFULL	-0.168	-0.05345
1581452	1866	LIFULL H	LIFULL	-0.10142	0.001344
159772	222	ニッセン	NITSE	-0.122	-0.122

CM数 : **168**
商品数 : 58

商品認知度



購買意向



■ : 広告効果あり(推定量>0)

(一部抜粋)

動画データの収集と前処理

<CM動画データの処理フロー>

動画取得

CM出稿データのcm_code・放送日・商品名から実際の動画を特定し、公式サイト/公式Youtubeからダウンロード[*]

前処理

各動画ファイルをmp4形式に変換し、秒数とフレームレートを統一
秒数：15秒，フレームレート：30fps [**]

[*] 分析対象の168CM中、**92CMの動画**が収集できた。

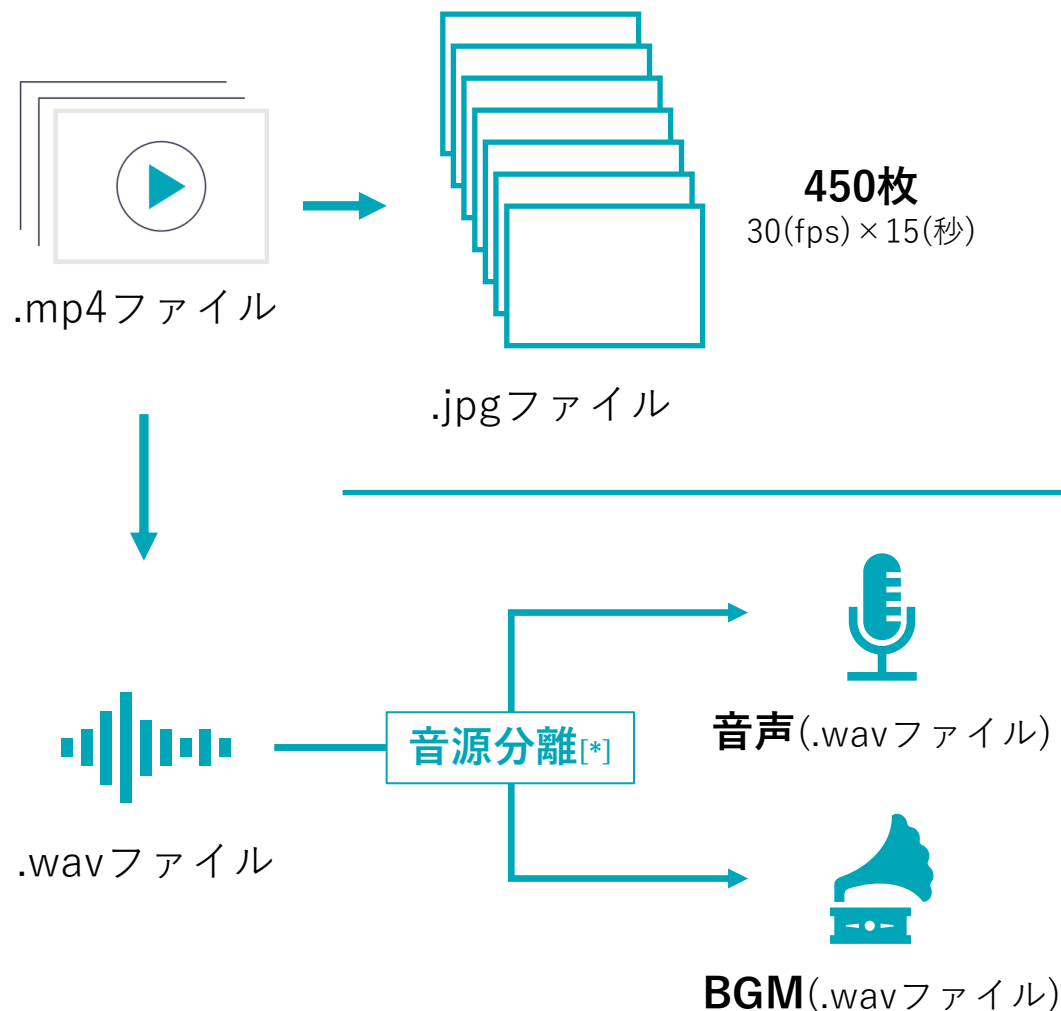
[**] フレームレートとは、1秒間の動画が何枚の画像で構成されているかを示すの単位。

fps：frames per second

基礎分析②

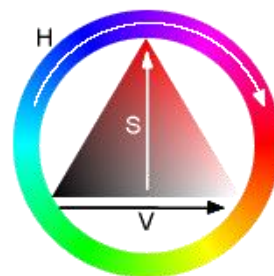
本分析

特徴量の抽出



HSV

色を色相/彩度/明度で表現したもの。



色相：180段階

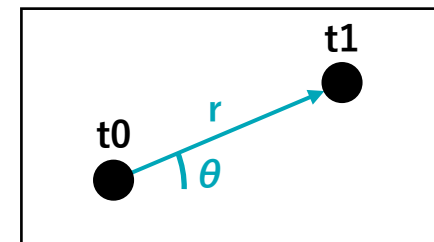
彩度：256段階

明度：256段階

(※ OpenCVでの表現)

Optical Flow

隣接するフレームで同一の画素を検出し、その移動の大きさと角度を算出する。



F0

F0(基本周波数)とは、音声の周期性を表現した、**ピッチ**を表す音響特徴量。(単位:Hz)

ex)

男性の声: Avg 125Hz, sd 20.5Hz

女性の声: Avg 250Hz, sd 41Hz [9]

LUFS

LUFS(**L**oudness **U**nit **F**ull **S**cale)とは、音声の**音圧**を測定する単位。

人間の聴覚特性に合わせて作られた規格で、実際に人が感じる**音量感**を表す。

[*] Deezer社(フランスの音楽配信サービスを行う)が開発したオープンソース「**spleeter**」を使用

本分析

提言

本分析の概要

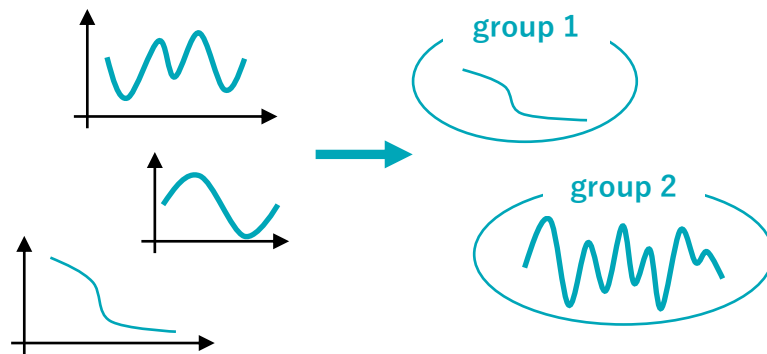
本分析① 視聴覚変数の作成

本分析② モデリング

背景

視聴覚特徴量(時系列データ)は高次元データであるため、モデリングの精度と解釈性を担保するのが難しい。

アイデア

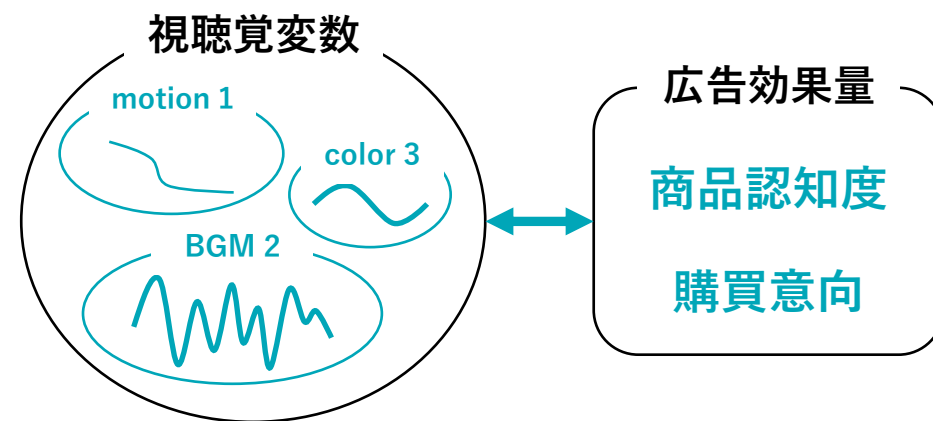


Fuzzy c-shape

クラスタ数：2~3 [*]

手法

- ・ 視聴覚変数をもとに広告効果量のモデリングをしたい。
- ・ 目的変数間の関係を捉える。



正準相関分析

[*]クラスタ数は各クラスタに属するデータ数の均衡性と、クラスタ重心の解釈性により決定

本分析

提言

Fuzzy c-shapeとは

時系列データのクラスタリング

時系列データは振幅や周期, 位相シフト, 垂直シフト等が変動する。
→ 拡大縮小や上下左右シフトに影響されない距離尺度が必要。 [1]

形状ベース(Shape based)のクラスタリング

距離関数: $SBD(\vec{x}, \vec{y}) = 1 - \max_w \left(\frac{CC_w(\vec{x}, \vec{y})}{\sqrt{R_0(\vec{x}, \vec{x}) \cdot R_0(\vec{y}, \vec{y})}} \right)$

重心ベクトル: 最小二乗距離

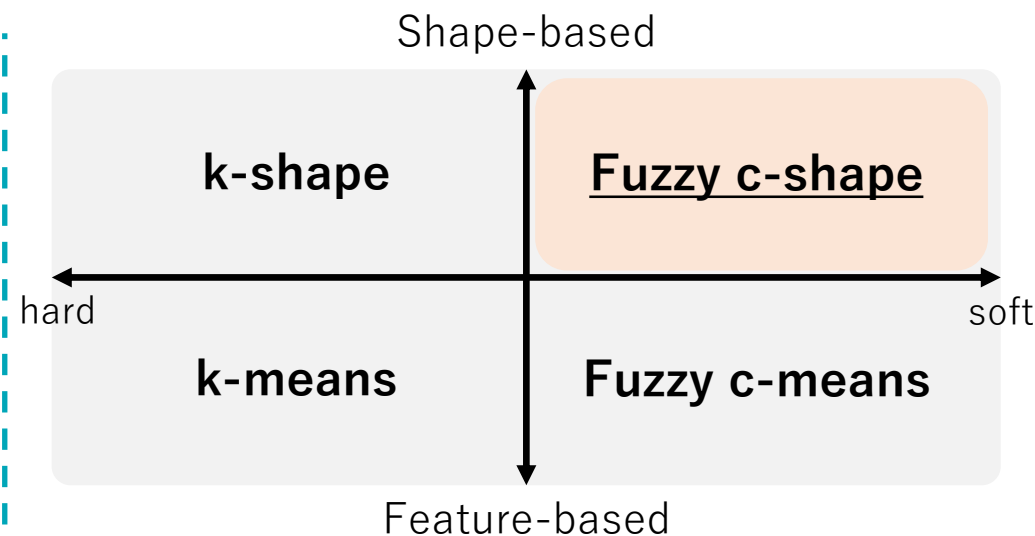
スケール変換やシフト変換によって変化しない距離尺度を用いるため, **時系列データの形状を保ったまま**比較できる。 [1]

ソフトクラスタリング

各データが**所属確率**に基づき, **複数のクラスタ**に属する。
(※ハードクラスタリング: 各データが1つのクラスタに属する。)

Fahimanら(2016)[2]は, k-shape[5]より

Fuzzy c-shapeの方が分類精度が優れていることを示した。



$CC_w(\vec{x}, \vec{y})$: 相互相関

$\sqrt{R_0(\vec{x}, \vec{x})}$: 自己相関の幾何平均

$\frac{CC_w(\vec{x}, \vec{y})}{\sqrt{R_0(\vec{x}, \vec{x}) \cdot R_0(\vec{y}, \vec{y})}}$: 正規化された相互相関

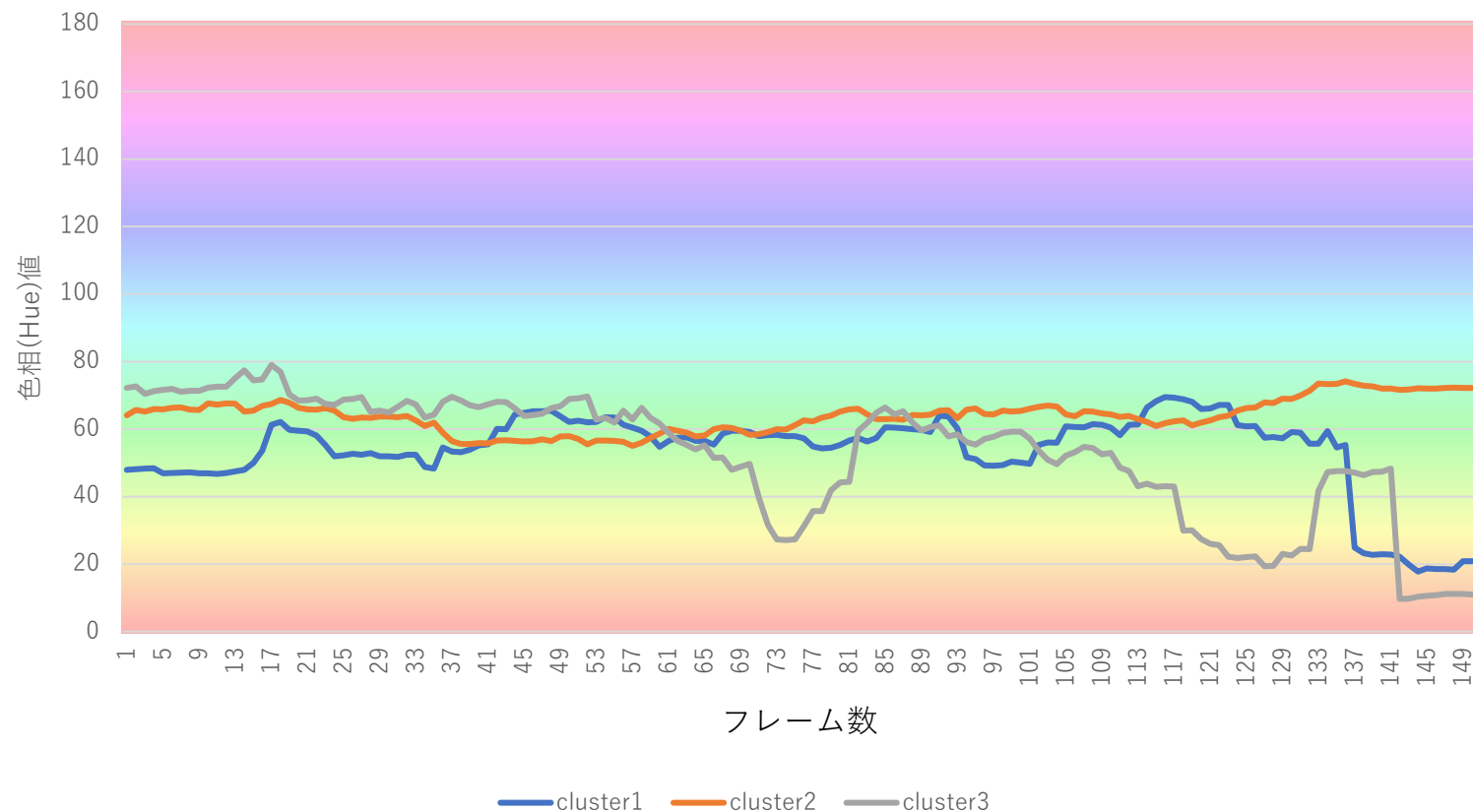
$SBD \in [0, 2]$: 0に近いほど類似

本分析

提言

クラスタリング結果①

color_h(色相)

**cluster1**

緑色を基調とし、動画の終わりで黄色やオレンジ色の表現が入る。

cluster2

緑色を基調とし、色の変動が小さい。

cluster3

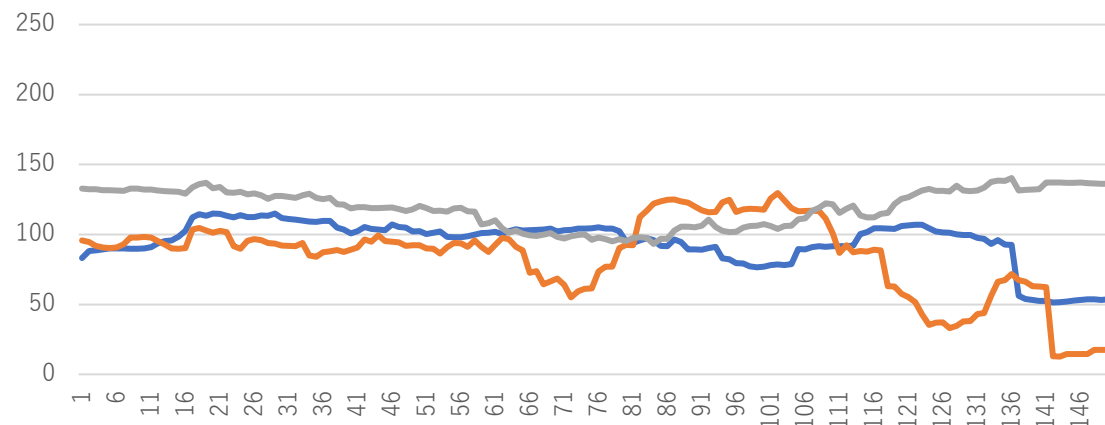
水色や緑色が基調で、中盤と後半で黄色やオレンジ色の表現が入る。

本分析

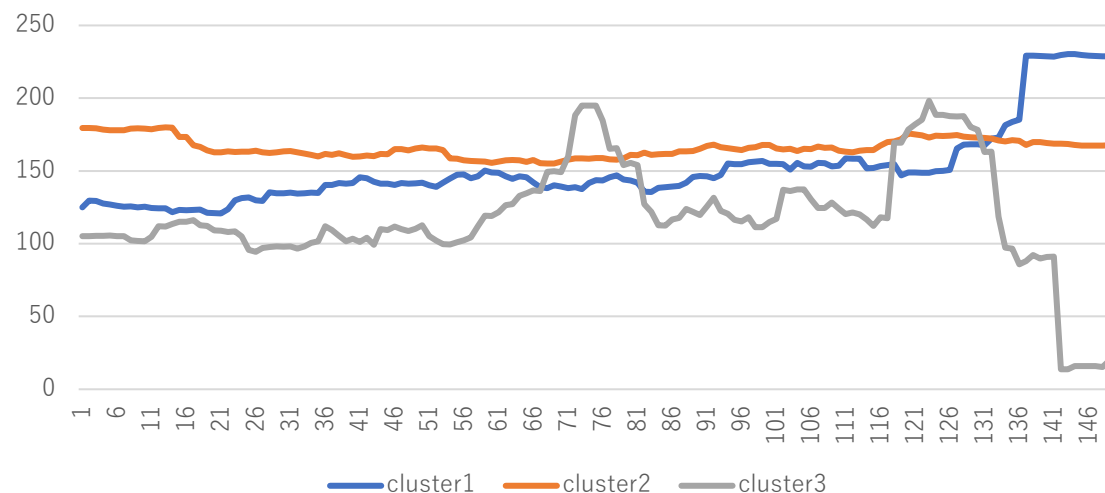
提言

クラスタリング結果②

color_s(彩度)



color_v(明度)

**cluster1**

彩度の変化が小さく，動画の終わりで彩度が落ちる．

cluster2

後半で彩度の変化が大きくなり，終盤は彩度が落ちる．

cluster3

彩度の変化は小さいが，前半は徐々に小さくなり，後半は徐々に大きくなる．

cluster1

明度が徐々に上がる．

cluster2

明度が一定で変化が小さい．

cluster3

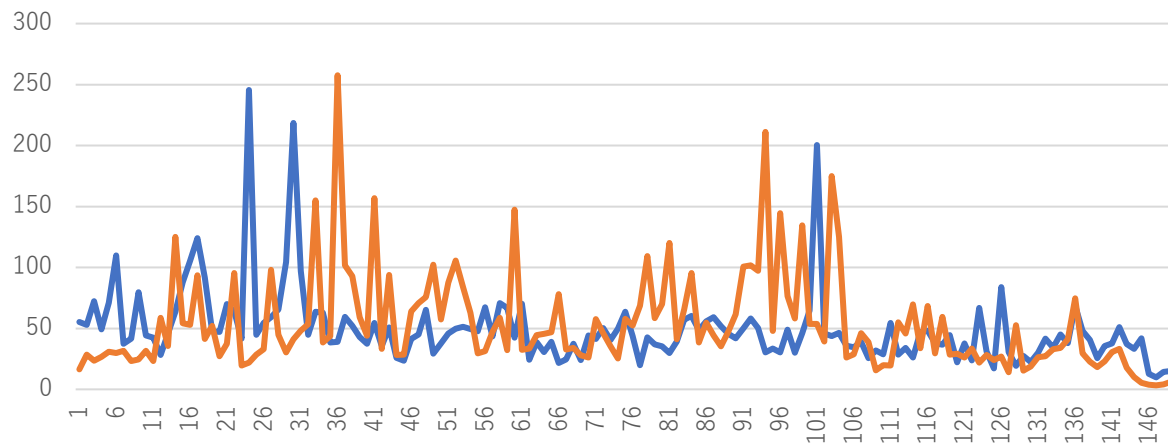
動画中盤で明度が大きくなり，動画の終わりでは急激に小さくなる．

本分析

提言

クラスタリング結果③

motion_vel_var(動きの大きさ(分散))



動きの大きさの**ばらつきが大きくなる**.
解釈：ある**動きを強調**する表現.

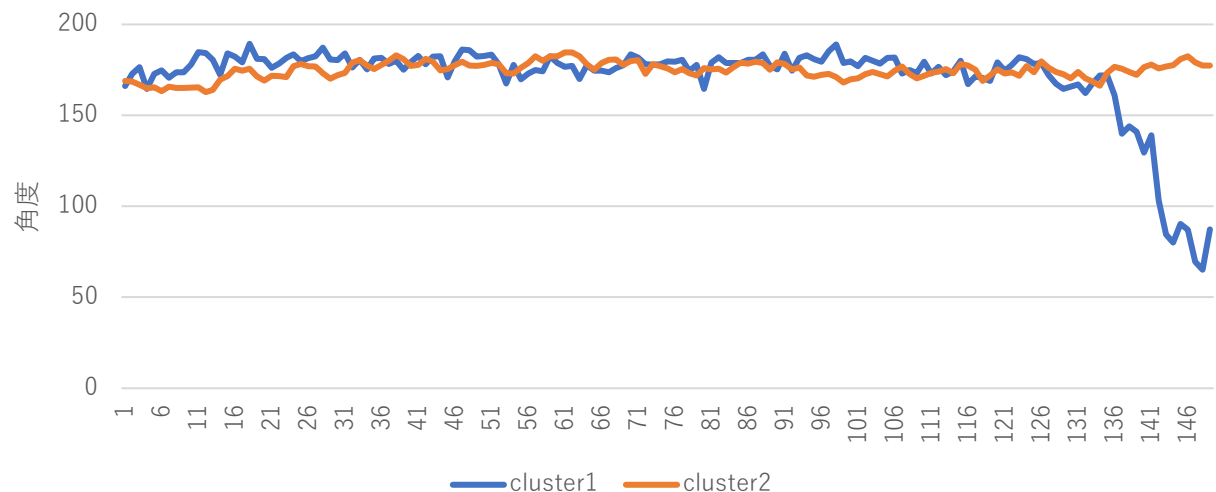
cluster1

比較的動きを強調する回数が少ない.

cluster2

動画中盤で特定の動きを強調する表現が増える.

motion_ang_mean(動きの角度(平均))

**cluster1**

基本的には左向きの動きが中心.

cluster2

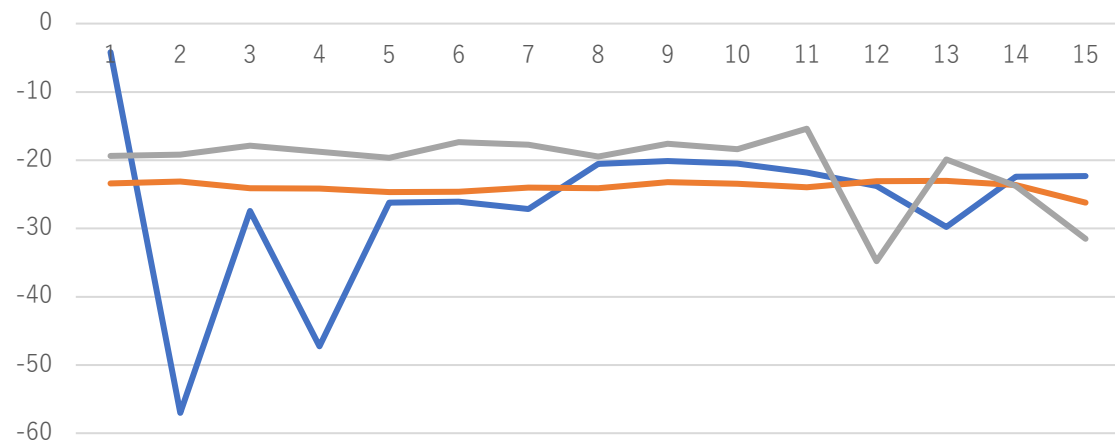
基本的には左向きの動きが中心で、動画終盤で左上～上方向の動きが増える.

本分析

提言

クラスタリング結果④

vocal_lufs(有声音の音量)

**cluster1**

動画始まりが最大音量で，前半に音量の変化が大きい。

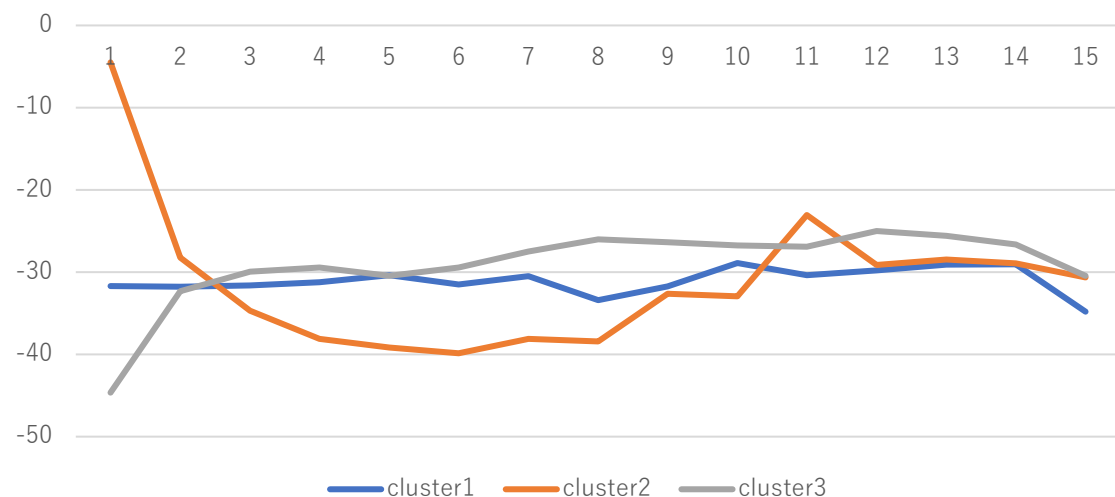
cluster2

音量の変化が小さい。

cluster3

動画終盤で二段階に分けて音量が小さくなる。

BGM_lufs(BGMの音量)

**cluster1**

動画始まりの最大音量から徐々に音量は小さくなり，後半で再度音量が大きくなる。

cluster2

音量の変化が小さい。

cluster3

動画の始まりは音量を抑えて，徐々に音量を大きくしていく。

本分析

提言

正準相関分析に用いる視聴覚変数群

変数名	内容	パラメータ
color_h_(i)	フレーム毎の 色相平均 値(各クラスタへの所属確率)	i= 1, 2, 3 (クラスタ名)
color_s_(i)	フレーム毎の 彩度平均 値(各クラスタへの所属確率)	i= 1, 2, 3 (クラスタ名)
color_v_(i)	フレーム毎の 明度平均 値(各クラスタへの所属確率)	i= 1, 2, 3 (クラスタ名)
motion_vel_mean_(i)	フレーム毎の 動きの大きさ平均 (各クラスタへの所属確率)	i= 1, 2 (クラスタ名)
motion_vel_var_(i)	フレーム毎の 動きの大きさ分散 (各クラスタへの所属確率)	i= 1, 2 (クラスタ名)
motion_ang_mean_(i)	フレーム毎の 動きの角度平均 (各クラスタへの所属確率)	i= 1, 2 (クラスタ名)
motion_ang_var_(i)	フレーム毎の 動きの角度分散 (各クラスタへの所属確率)	i= 1, 2 (クラスタ名)
vocal_f0_(j)	有声音の周波数	j=[" mean ", " min ", " max "]
BGM_f0_(j)	BGMの周波数	j=[" mean ", " min ", " max "]
vocal_lufs_(i)	1秒ごとの 有声音ラウドネス値 (各クラスタへの所属確率)	i= 1, 2, 3 (クラスタ名)
BGM_lufs_(i)	1秒ごとの BGMラウドネス値 (各クラスタへの所属確率)	i= 1, 2, 3 (クラスタ名)

本分析

提言

正準相関分析とは

正準相関分析は、2つの多次元連続値確率変数について線形射影を行うことで、高い相関を持つ線形部分空間の対[*]を見出す手法である。[10]

x_i : 視聴覚特徴量に関する変数($i = 1, 2, \dots, 25$)
 y_1, y_2 : 広告効果量に関する2変数(商品認知と購買意向)
としたとき,

$$X = \alpha_1 * x_1 + \alpha_2 * x_2 + \dots + \alpha_{25} * x_{25}$$

$$Y = \beta_1 * y_1 + \beta_2 * y_2$$

X, Y の相関係数を最大化するような, α, β を求める。
この X, Y を**正準変量**, α, β を**正準変量係数**という。

[*]正準変量が少ない方の数量分(今回の場合は2個)の結果が出る。

本研究への適用

得られた視聴覚変数群の
前処理。(後述)

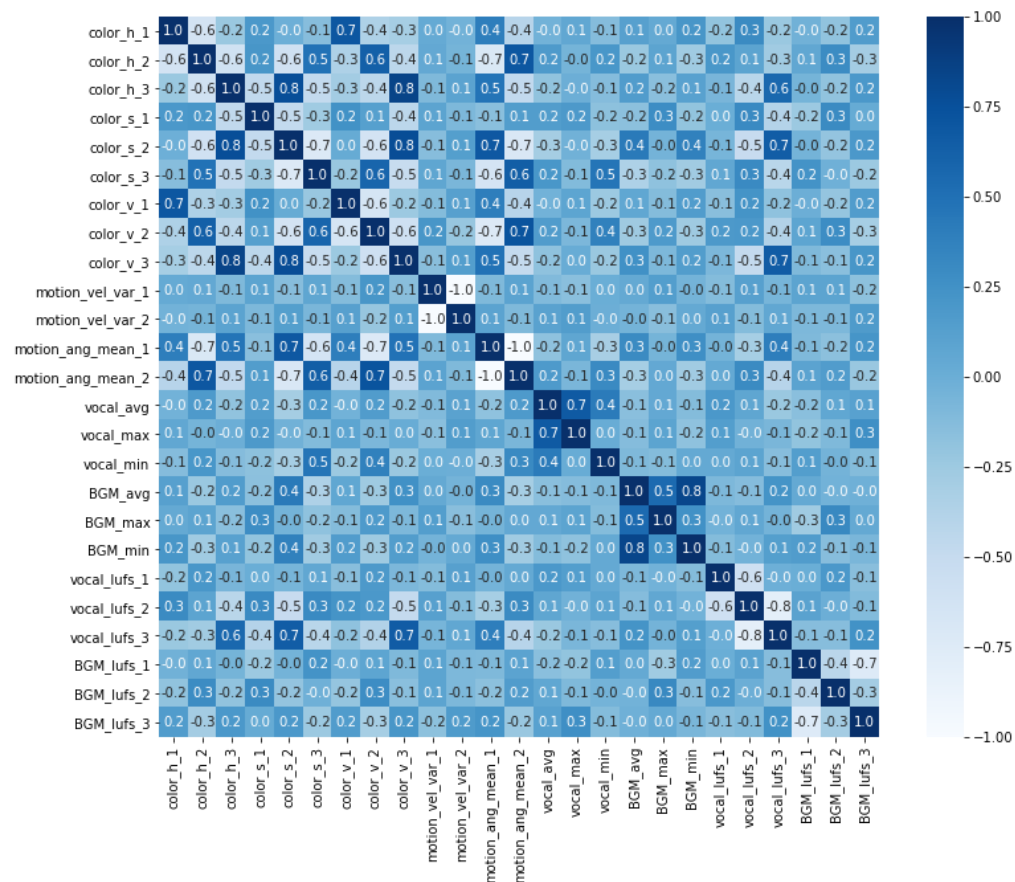
正準相関分析

正準相関と**正準変量係数**をもと
に**広告効果**と**視聴覚変数**の関係を
可視化する。

視聴覚変数群の前処理

変数選択

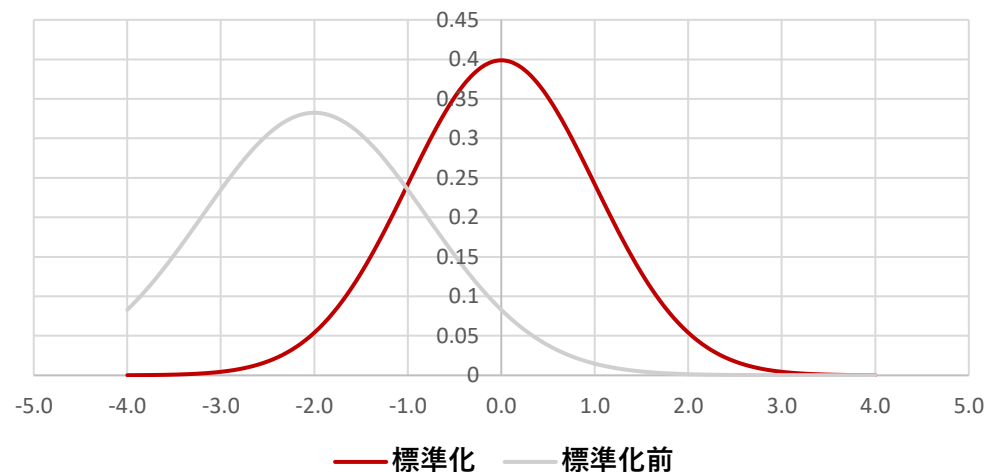
視聴覚変数間で相関の高いものを
正準変量から除外。（目安： $|p| > 0.7$ ）



標準化

分析結果を正準変量係数で解釈するため、
分布を統一する。

標準化イメージ

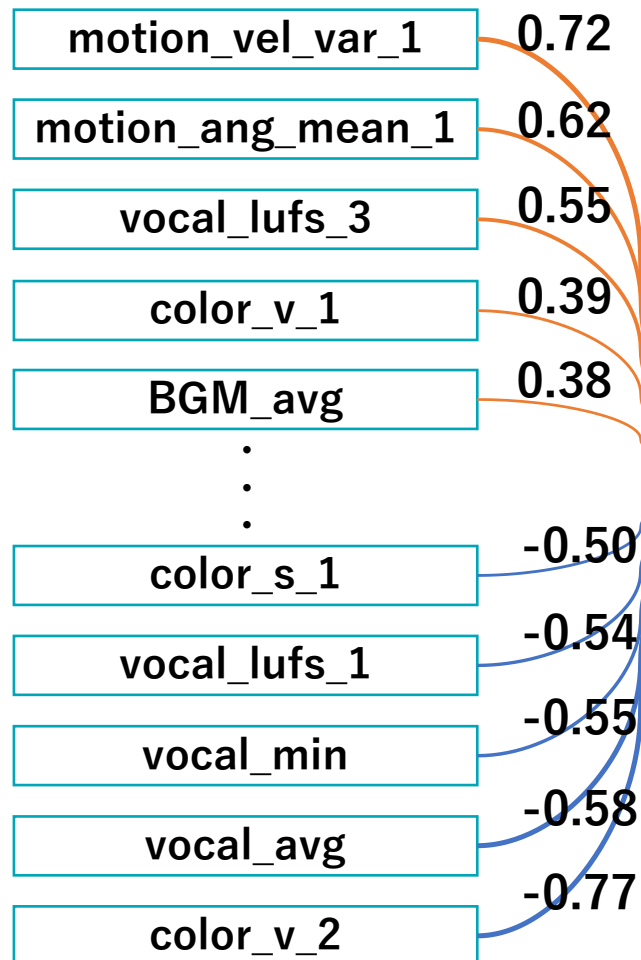


本分析

提言

正準相関分析の結果

結果①



商品認知

購買意向

2.08

0.86

1.53

0.51

広告効果①

広告効果②

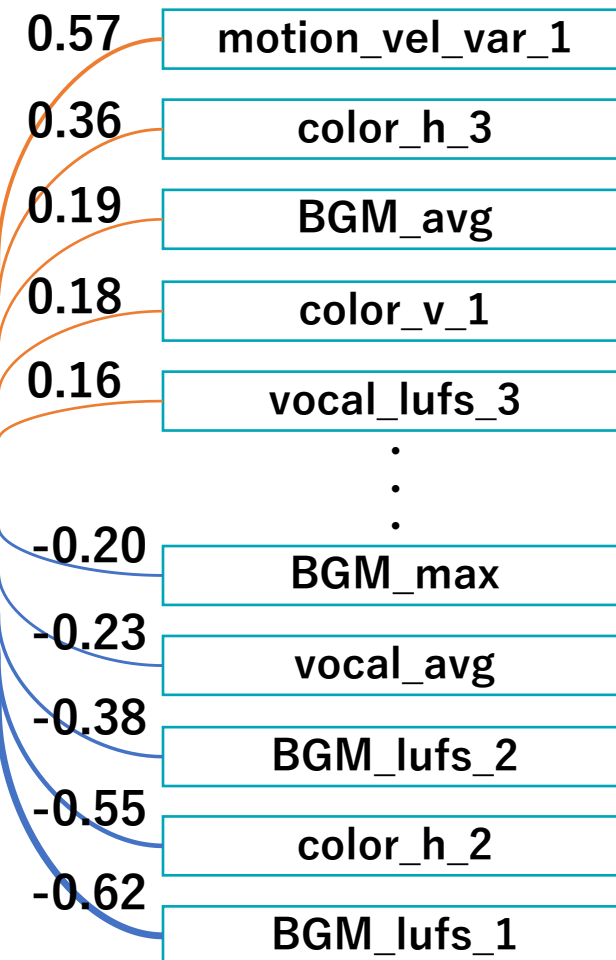
0.474

0.404

視聴覚特徴①

視聴覚特徴②

結果②



正準変量係数が上位5件・下位5件の変数を記載

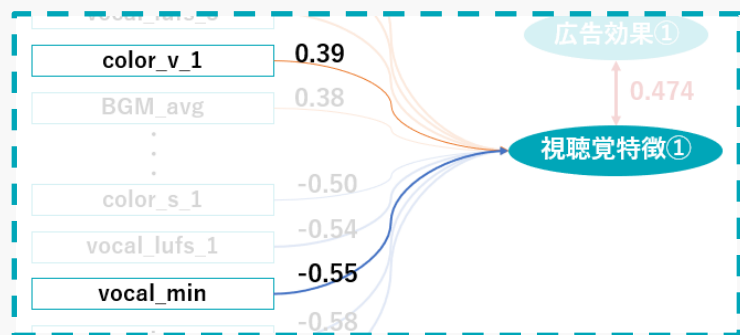
本分析

提言

正準相関分析の解釈例

ex1) 正準変量係数による正準変量 Y の解釈

商品認知度と購買意向がともに高まるような広告効果。
特に、**商品認知度**の効果が高くなる。

ex2) 正準変量係数による正準変量 X の解釈

色の明度クラスター1への所属確率を表す係数が**正**である。
→明度特徴：中盤で大きくなり，終盤で急激に小さくなる。
有声音の周波数最小値を表す係数が**負**である。
→音程特徴：比較的**男性音声**が使われている特徴[*]

ex3) 正準変量 X, Y の正準相関の解釈

視聴覚特徴と広告効果に正の相関がある。
→**視聴覚特徴**が大きいと**広告効果**も大きい傾向がある

[*]一般的に男性の方が女性より，音声の周波数が低いため。

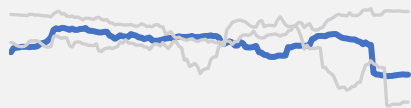
本分析

結果の解釈

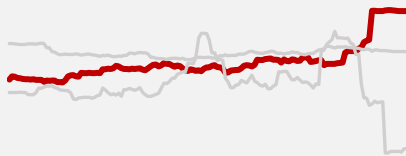
提言

<特に商品認知度と関連が強い特徴>

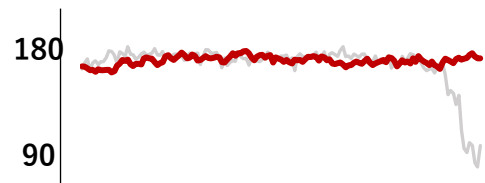
彩度 中盤での彩度に変化を加える。



明度 徐々に大きくなる。

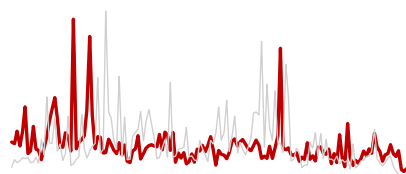


色

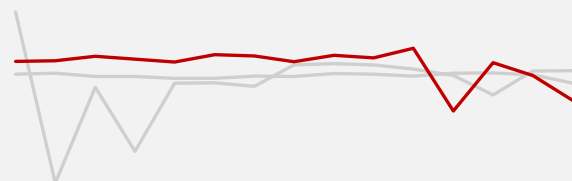


角度(平均) 一貫して左向きが中心。

大きさ(分散) [*]と同様



動き



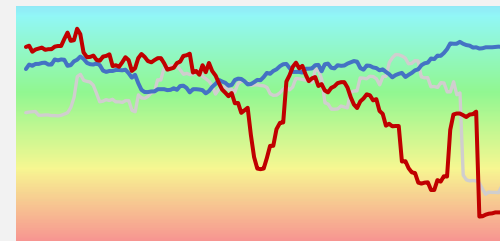
vocal 比較的大きく、動画終盤で二段階で小さくなる。[**]

音量

音程

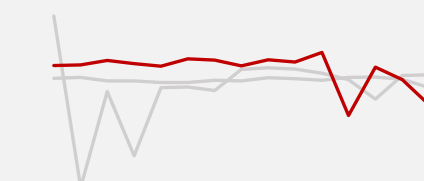
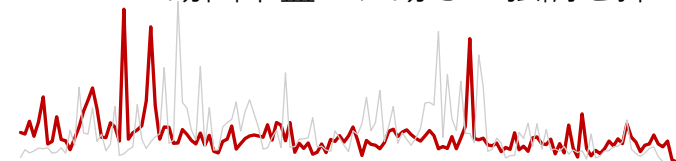
比較的BGMの音が高い。
男性音声が多めに含まれる。

<特に購買意向と関連が強い特徴>

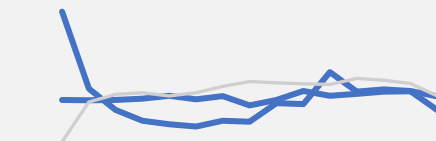


色相
水～緑色が基調。
中盤と終わりで黄～赤色の表現が入る。
色相に変化をつける。

大きさ(分散) 特定の動きを強調しない。
動画中盤では動きの強調を抑える。[*]



vocal [**]と同様



BGM 始まりは抑えて、
徐々に大きくなる。

比較的BGMの音が高い。
男性音声が多めに含まれる。

提言

本研究の展望

・ 広告クリエイティブ戦略への寄与

認知度の拡充に重点を置く場合



状況例：新商品のプロモーション

本分析の結果①に基づく
クリエイティブ案の考察

購買意向の拡充に重点を置く場合



状況例：既存広告の見直し

本分析の結果②に基づく
クリエイティブ案の考察・ データ拡充への対応

本研究で用いた視聴覚特徴は全てCMの動画データから機械的に抽出できる物理量である。
そのため、CMデータの追加やその量に関わらず特徴量を抽出し、広告効果の検証に使用することができる。



新規動画データ

Get Visual/Audio Cue

本研究で作成した
特徴量抽出パイプライン

応用例

動画の視聴覚特徴を可視化
推定モデルへの発展

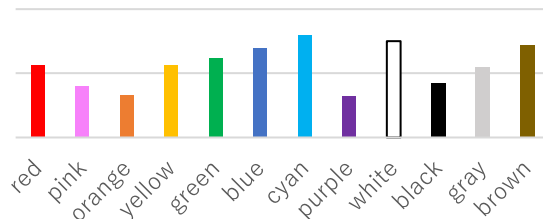
提言

本研究の課題

・特徴量抽出の精緻化

本研究での視聴覚特徴量は画像全体の色相平均値など、大まかな統計量として扱った。
下記のような抽出処理の改善で、視聴覚特徴の情報損失が小さくなり、**解釈性の向上**を期待できる。

色特徴



色彩心理に基づく解釈

動き特徴



<ツール例>
OpenCV
Cloud Vision API

人の**出演/動作**を
特徴量化

有声音特徴



<ツール例>
inaSpeechSegmenter
音声特徴を**性別**で分割

テキスト情報の取得



「あいうえお」



<ツール例>
inaSpeechSegmenter
Speech Recognition
Cloud Vision API

・購買意向の先の消費行動まで捉えた分析

本研究では、広告効果として商品認知度と購買意向の2種類を用いた。
これらの指標に加えて**購買行動**の情報を追加することで、**広告の視聴覚特徴と視聴者の購買プロセス**の関係を探ることができる。

購買プロセスの参考) AIDA[6] **Attention** → **Interest** → **Desire** → **Action**

- [1] Ankur, A. P. (2019). Hands-On Unsupervised Learning Using Python: How to Build Applied Machine Learning Solutions from Unlabeled Data. O'Reilly & Associates INC. (Ankur, A. P. 中田秀基 (訳). (2020). Pythonではじめる教師なし学習 機械学習の可能性を広げるラベルなしデータの利用. オライリージャパン)
- [2] Fahiman, F., Bezdek, J. C., Erfani, S. M., Palaniswami, M., & Leckie, C. (2017). Fuzzy c-Shape: A new algorithm for clustering finite time series waveforms. In 2017 IEEE International Conference on Fuzzy System(FUZZ-IEEE) (pp. 1-8). IEEE.
- [3] Krishna, A. (2012). An integrative review of sensory marketing: Engaging the senses to affect perception, judgment and behavior. Journal of consumer psychology, 22(3), 332-351.
- [4] Krishna, A., Cian, L. & Sokolova, T. (2016). The power of sensory marketing in advertising. Current Opinion in Psychology, 10, 142-147.
- [5] Paparrizos, J. & Gravano, L. (2015). k-shape: Efficient and accurate clustering of time series. In Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data(pp. 1855-1870)
- [6] Strong Jr, E. K. (1925). Theories of selling. Journal of applied psychology, 9(1), 75.
- [7] Tao, L., Wang, X., Kawahara, T. & Yamasaki, T. (2020). Television Advertisement Analysis Using Attention-based Multimodal Network. In 人工知能学会全国大会論文集 第34 回全国大会, 1H4OS12b01-1H4OS12b01.
- [8] 赤井良行, 李昇姫(2014). 音色からイメージされる色彩の歓談と音色構造の関係. 日本感性工学学会論文誌, 13(1), 221-228
- [9] 一般社団法人日本音響学会. “音のなんでもコーナー 050”, <https://acoustics.jp/qanda/answer/50.html>.(2021/10/20 アクセス)
- [10] 岩瀬智亮, 中山英樹(2016). 深層一般化正準相関分析. 第 78 回全国大会講演論文集, 2016(1), 183-184.
- [11] 永富美里, 森達平, 中江菜々子(2019). 『「音」から見るテレビCMの広告効果の分析』.
<https://www.is.nri.co.jp/contest/2019/download/mac2019saiyushu.pdf>.(2021/07/13 アクセス)
- [12] 星野崇宏(2009). 調査観察データの統計科学 因果推論・選択バイアス・データ融合. 岩波書店.
- [13] 延岡健太郎(2010). オープン・イノベーションの陥穽: 価値づくりにおける問題点(特集「オープン・イノベーション」の再検討). 研究 技術 計画, 25(1), 68-77.
- [14] 延岡健太郎, 伊藤宗彦, 森田弘一(2006). コモディティ化による価値獲得の失敗: デジタル家電の事例. イノベーションと競争優位-コモディティ化するデジタル機器-』NTT 出版, 14-48.
- [15] 清水柚里奈, 菅野沙也, 伊藤貴之, 嵯峨山茂樹, 高塚正浩(2016) 動画特徴量からの印象推定に基づく動画 BGMの自動生成. 情報処理学会第 78 回全国大会, 2Q-01.
- [16] 白石弘幸(2013). コモディティ化と製品の心理的価値. 金沢大学経済論集, 34(1), 113-154.