

映像のシーンに対応した視聴者の感情・印象抽出

山口 想^{1,a)} 梁木 俊冴³ 高屋 英知³ 高野 雅典² 森下 壮一郎² 山本 悠二² 福田 鉄也²
福田 一郎² 栗原 聡³

概要：番組制作者は、視聴率や再生数を上げようと努力する。視聴率を上げるためには視聴離脱を防ぐことが不可欠である。視聴離脱の要因の一つに視聴者の映像コンテンツに抱いた感情や印象などの捉え方があると考えられる。しかし、映像コンテンツから直接、視聴者の捉え方を抽出することは困難である。そこで番組実況という視聴者が映像を視聴しながら書き込むことができるコメントに着目した。番組実況に含まれる視聴者の捉え方をトピックモデルや文字レベル畳み込みニューラルネットを用いて抽出を試みた。そして抽出した視聴者の捉え方を要因として視聴離脱の要因分析を行った。

Extract impressions of viewers corresponding to scenes of video

1. はじめに

近年、映像の配信プラットフォームはテレビだけに留まらず、パソコンやスマートフォンなど、多岐にわたっている。これらのデバイスを用いることで、老若男女問わずいつでもどこでも映像の視聴を行うことが可能となっている。マクロミルの調査によると、日本における映像コンテンツの利用率は98%とほぼ全ての人が利用していると言える[1]。また、総務省情報通信政策所の調査によると10代から20代の若者はテレビよりネット利用時間の方が長く、YouTubeの使用率は90%という結果となった[7]。そしてNHK技研の報告書によるとインターネット動画の利用実態によると20代の90%がYouTubeを利用し、50%がニコニコ動画^{*2}を利用することがわかった[11]。インターネット動画を視聴できるサービスとしては、上述したYouTubeやニコニコ動画といった動画共有サービスに加え、リアルタイムで配信が行われるTwitch^{*3}やAbemaTV^{*4}のようなライブストリーミングサービスと呼ばれるものなどが存在し

ている。これらの動画共有サービスやライブストリーミングサービスの多くは、広告収入によって運営されており、基本的に無料で利用することができる。また、国内における民放局も広告収入を主な収入源としており、各サービスの広告枠を企業に買ってもらうような形をとっている。人気の動画や視聴率の高い番組、つまり多くの人が視聴する映像ほど広告の宣伝効果が高いため、番組制作側は、視聴率や再生数を上げようと努力する。ここで視聴率とは、テレビ番組や広告がどのくらいの世帯や人々に見られているかを示す指標の一つとしてテレビ局などで導入されている。日本では現在、ビデオリサーチ社が測定を行なっている。ビデオリサーチ社によると視聴率データは「国民の関心の高さを知る」「社会の動きを知る」という社会調査的な側面での利用、また広告出稿者、テレビ局、広告会社が広告取引をする際にも、番組の媒体力や広告効果を測る一つの指標として利用されている。つまり、番組の配信プラットフォームやテレビ局はいかにして番組の視聴率を上げるか考える必要がある。特に、視聴率を上げるためには視聴離脱を防ぐことが不可欠である。視聴離脱とは、番組視聴途中に視聴を中断することである。視聴者が視聴離脱する要因は様々であるが、その一つに視聴者の映像コンテンツに抱いた感情や印象などの「捉え方」があるだろう。例えば、視聴していて不快な内容や視聴者にとって嫌いなキャストが登場した際に視聴離脱が起こると考えられる。そこで、番組コンテンツにおける番組シーンごとの視聴者の「捉え方」を抽出し、視聴離脱の原因を説明できるか検討を行う。

¹ 電気通信大学 情報理工学研究所
The University of Electro-Communications, Chofu, Tokyo 182-9585, Japan
² 株式会社サイバーエージェント
CyberAgent, Inc., Shibuya, Tokyo 150-0043, Japan
³ 慶應義塾大学 理工学研究所
Keio University, Minato, Tokyo 108-8345, Japan
^{a)} y1730110@edu.uec.ac.jp
^{*2} ニコニコ動画 <https://nicovideo.jp>
^{*3} Twitch.tv <https://www.twitch.tv/>
^{*4} AbemaTV <https://www.abema.tv/>

しかしながら、映像コンテンツから直接、視聴者の感情や印象などの「捉え方」を抽出することは困難である。そこで映像を視聴した視聴者の間接的な情報を用いて「捉え方」の取得を試みた。従来研究では、間接的な情報を用いて視聴者の映像コンテンツからの印象を取得する手法としてアンケート評価や脳活動計測が存在する。金ら [2] は、『映像を見る人によって異なる見方があることを理解するメディアリテラシーの向上を支援するため、映像に対する視聴者の反応を把握する実験』を行なった。彼らは被験者に映像を視聴させ、その後にアンケートによる印象評価を行なった。だが、アンケート評価も脳活動計測も多数の番組についてた多人数に行ってもらい、つまり大規模に行うとコストがかかってしまう。そこで本研究では映像視聴時の番組実況に着目した。

番組実況とは、視聴者がスポーツ中継や映画などをテレビで視聴しつつ、その時々の内容や感想、意見などをインターネット上のサービスにリアルタイムに書き込み、会話を楽しむインターネット文化である。利用者は、映像コンテンツを視聴しながら並行してチャットを楽しむことにより、視聴者との感情の共有による一体感を味わうことができる。この番組実況を用いて視聴者の「捉え方」の抽出を行う。

本研究では、各番組における番組実況に含まれる視聴者の「捉え方」が視聴者全体の意見や感想であると仮定する。そして多くの視聴者が視聴離脱するシーンにおける番組実況を分析し、番組のシーンにおける離脱傾向を理解することを目的とする。

2. 番組実況

番組実況に関する研究は、2000年代前半から書き込む媒体がインターネット掲示板、Twitter と対象を変え行われてきた。宮森ら [9] はインターネット掲示板「2ちゃんねる」における番組実況スレッドの書き込みを用いて、インターネット特有の汎用的な特徴に着目した。そして、番組実況は番組の盛り上がりや嗜好・興味に沿った反応を効率良く抽出できる可能性が高いことが判明した。山内ら [8] は、番組放映時の Twitter の各シーンごとのツイートに対して感情分析を行い、視聴者側の意見に基づくインデキシングを試みた。近年では、リアルタイムに番組を視聴して、実況を行うのではなく、動画に擬似的に実況を行うことも注目されている。日本のニコニコ動画や中国の Bilibili などの動画共有サイトにおいて、投稿された動画は再生時間軸上に対し視聴者がコメントを投稿できる特徴を持っているため、そのコメントデータを用いることによって従来の番組実況と類似したことができる。番組実況と擬似実況の共通の問題として、番組の視聴から書き込みまでどうしても遅延が発生してしまう問題がある。Qing ら [10] は、視聴からコメント投稿までの遅延を考慮するため、投稿された

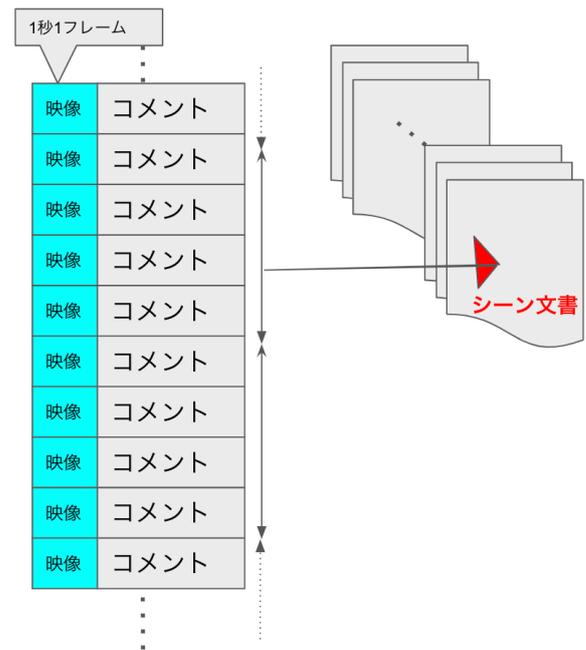


図1 番組の定数幅で分割し各区間からシーン文書を作成する流れ

コメントをラベリングして一定の区間における同じラベルが付与されたコメントに対して遅延の補正を行った。そして、遅延補正済みのコメントデータを用いて、動画像を見ずに動画に対するハイライト生成を行った。

番組の映像と番組実況に関する研究や番組実況を用いた映像に対するメタデータの研究は多く行われているが、番組実況と視聴者の視聴行動に関する研究は見当たらなかった。

3. データセット

本研究では、AbemaTVにてリアルタイムに放映された「恋愛リアリティ」というジャンルに属する番組3作品60話分の視聴者の投稿コメントおよび放映時の視聴者の視聴ログを用いた。恋愛リアリティというジャンルを対象にした理由は以下の通りである。

- AbemaTVにて人気コンテンツなため視聴者数とコメント数が多い
- 視聴者の投稿するコメントが感情や印象といった内容を多く含めている

また、各番組のシーンについて脚本などのメタデータを利用して明示的に分割することが理想であるが、現実的にはそのようなメタデータの入手は困難である。そのため、本実験では定数幅で分割を行った(図1)。分割した各区間を1シーンと定義し、その区間内に投稿されたコメントがそのシーンと対応させた。「恋愛リアリティ」というジャンルでは番組の展開が時時刻々と変化するため、10秒が適任であると判断した。

4. 各印象トピックの視聴離脱の影響度合い

本節では、どのような印象が視聴離脱に影響するのかについて分析を行った。

印象トピックの抽出

まず、各シーンに対応したコメントを結合することによってシーン文書を作成する(図2上)。作成したシーン文書に対して、繰り返し正規化を行った後、分かち書きを行い、品詞が形容詞の単語のみで再構成したシーン文書を作成する。トピックモデルの一種であるディリクレ配分法(Latent Dirichlet Allocation, LDA)[4]を用いて、作成したシーン文書からシーンの印象トピックを抽出する。本実験では抽出するトピック数を30とした。印象トピックの命名については、番組のジャンルによって同じ形容詞でも解釈が異なることも考慮し、構成する形容詞およびトピックのスコアが高いシーンの映像を視聴することによって命名した(表1)。

視聴離脱に影響する印象の分析

シーンごとのシーン文書からシーンを構成する印象トピックの抽出を行う(図2下)。そして、どの印象トピックが視聴離脱に影響しているか知るために、先ほど算出したシーンの各印象スコアを説明変数、シーンの視聴離脱数を目的変数とする回帰モデルを構築した(数式1)。回帰モデルは、視聴離脱数がポアソン分布に従うと仮定するポアソン回帰の一般化線形モデルにパラメータ推定と変数選択を同時に行うため正則化項としてL1ノルム[6]を追加した正規化回帰モデルとした。交差検証を行い、最適なハイパーパラメータ λ を求め、その λ における各回帰係数を求めた。

$$\hat{\beta}_{Lasso} = \underset{\beta}{\operatorname{argmin}} \left\{ \frac{1}{2} \sum_{n=1}^N \left(y_n - \beta_0 + \sum_{j=1}^p x_{nj} \beta_j \right)^2 + \lambda \sum_{j=1}^p |\beta_j| \right\} \quad (1)$$

結果

構築した回帰モデルの各印象トピックの係数は表4となった。結果から、印象トピックとして「ポジティブ」と「卑劣・非難」が視聴離脱に影響していることがわかった。

結果の解釈

ポジティブな要因が、視聴離脱に影響するのは直感に反していた。そのため、次の分析では視聴離脱に影響するポジティブ、ネガティブの割合のパターン分析を行った。

5. 視聴離脱に影響するパターン分析

各番組の視聴離脱の顕著なシーンのみを抜粋し、そのシーンにおける各コメントに対して、ポジティブ、ネガティブ、ニュートラルのいずれかのラベルの付与を行う感情極性分類を行った。その後、シーンを構成するポジティブ、ネガティブ、ニュートラルの割合を算出し、可視化した。

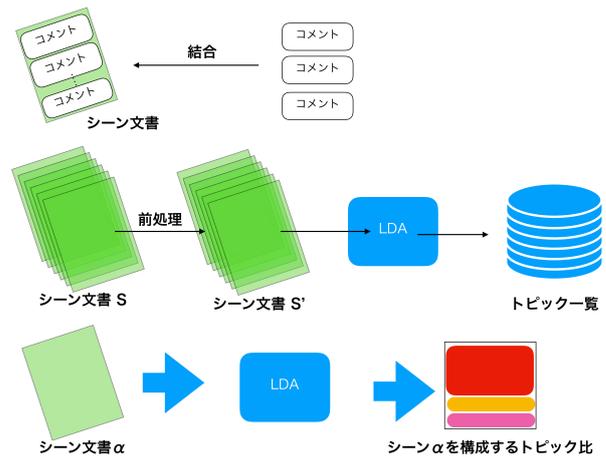


図2 上:コメントからシーン文書を作成する手順, 中:シーン文書からトピック一覧を取得する手順, 下:シーン文書からシーンを構成するトピックの割合を求める手順

番号	トピック名	構成する代表的な形容詞	係数
1	対立	好き, 可愛い, 良い, 酸い	0
2	ポジティブ	良い, 好きだ, 可愛い	0.0107
3	ポジティブ	旨い, 良い, 凄い	0
4	批判	速い, 無い, 不安だ	0.00210
5	-	嬉しい, 良い, やばい	0.00000632
6	和やか	可愛い, 良い, 好きだ	-0.0000487
7	対立	怖い, 良い, 可愛い, 悪い	0.000689
8	ポジティブ	優しい, 良い, 可愛い	0
9	-	熱い, 可愛い, 良い	0.00148
10	ネガティブ	うるさい, だるい, キモい	0.000272
11	-	古い, 良い, 楽しい	0
12	重々しい	やばい, やだ, 怖い, 気不味い	-0.000152
13	-	広大大, 健勝だ, のんきだ	-0.00105
14	不穏	やだ, 良い, 駄目だ	0.0000428
15		良い, 好きだ, 無い	0.000556
16	ネガティブ	汚い, プスだ, 変だ	-0.00285
17	楽しい	長い, 良い, いい, 優しい	-0.00196
18	対立	良い, 好きだ, 嫌いだ	0.00180
19	-	雄大大, 良い, 濃い	0
20	卑劣/非難	可哀想だ, 最悪だ, 無い, 辛い	0.000372
21	-	速い, 浮気だ, 可愛い	0.0000894
22	-	格好良い, 好きだ, 良い	0.000140
23	ポジティブ	好きだ, 可愛い, 遙かだ	0.00304
24	-	早い, 無い, やだ	0.00668
25	-	巧みだ, 良い, 好きだ	0.000936
26	-	良い, 雄大大, 欲しい	0.00187
27	静寂	重い, 良い, 無い	-0.000808
28	-	良い, 薄い, 臭い	0
29	ポジティブ	良い, 好きだ, 無い	0.000833
30	ポジティブ	良い, 可愛い, 最高だ	0.000982

表1 印象トピックおよび視聴離脱回帰モデルの各印象トピックの係数

結果

各シーンを構成するポジティブ、ネガティブ、ニュート

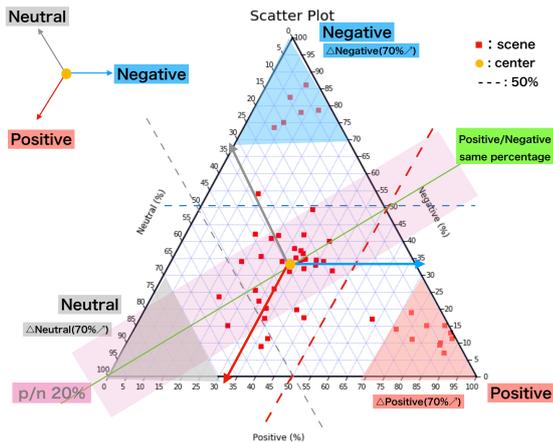


図3 視聴離脱の顕著なシーンにおけるポジティブ、ネガティブ、ニュートラルの割合を算出し、可視化した結果

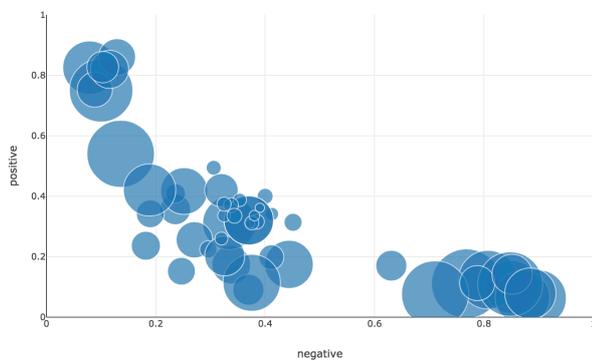


図4 シーンを構成するポジティブ、ネガティブの割合およびコメント数を可視化したバブルチャート

ラルの割合を算出し、可視化した (図3)。結果を見ると、基本的にはポジティブとネガティブの割合が均一かつ半数程度を占めている場合、また、ネガティブまたはポジティブが大多数の場合において視聴離脱が特に発生することがわかった。また、割合だけでなくシーン全体のコメント数を円の大きさで示したバブルチャートでも可視化するとコメント数が極端に多いときにおけるネガティブもしくはポジティブが大多数の場合に視聴離脱が発生することもグラフから読み取れた (図4)。

結果の解釈

ポジティブなコメントが多数みられる場合においても視聴離脱が発生しやすくなっている事実については一考の余地がある。考えられる理由としては、多数派であるポジティブなコメントに対し、少数派であるネガティブな意見を持つ人々がコメントを残す意欲を失い、離脱していることが挙げられる。「恋愛リアリティ」というジャンルでは、視聴者の多くが演者に対して感情移入しており、演者の行動やその結末に対し、共感や批判が発生すると考えられる。番組実況の場では、自分の感覚が大多数と異なっている場合、それが明確に可視化されるため、自らが少数派であるという事実はコメント意欲を喪失させる可能性がある。

label	ラベル	例
abusive	暴言	ks, 氏ね
advice	助言	~した方がよくな
critic	批判	ダメじゃね, おもんな
cry	感動	; ; , 泣いた
exciting	期待	ドキドキ, wtk
fear	恐怖	ぎゃあああ
feed	ツッコミ	なんでやねん
fever	熱狂	うおおおお, きたー!!!!
irritate	イライラ	しろよ, やれや, やめろ
like	賞賛	すげえ,
lol	笑い	w, 草, おもろ
question	疑問	?
request	要望	~して欲しい
response	返事	なるほど, ですよねー
suprise	驚愕	すごっ, すげえ,
greeting	挨拶	おはよう,
unknown	分類不能	上記に該当しない

表2 本分析で用いる感情・反応ラベル一覧

今までの分析にて、ポジティブ、ネガティブという要因と視聴離脱の関係性については判明した。次の分析では、インターネット用語に適応した、より粒度の細かいラベルをコメントに付与し、視聴離脱に影響する要因分析を行った。

6. 各感情や反応の視聴離脱の影響度合い

感情分類モデル

コメントデータに対して一言コメント感情分析 API^{*1}を用いてラベリングを行った。その後、ラベリングしたコメントデータを学習データとして、文字レベル畳み込みニューラルネットワーク (Character-Level Convolutional Neural Network, CLCNN) を用いて感情分類モデルを構築した。従来は、分かち書きを行い単語分割を行って入力としていた。しかし、番組実況のようなインターネット用語や崩れ文字、タイピングミスが含まれた不自然な日本語が多い場合において、文字単位で処理することはモデル自体が堅牢になり、また分かち書きする必要がないため最善だと思われる。また分析で用いた感情ラベルを表6に示す。

視聴離脱に影響する感情の分析

シーンごとの各コメントについて感情分類モデルを用いてラベリングを行う。ラベリングを行うことによって、シーンに含まれる各感情がどのくらい占めているか算出することができる。どの感情が視聴離脱に影響しているか知るために、先ほど算出したシーンの各感情数を説明変数、シーンの視聴離脱数を目的変数とする回帰モデルを構築し各回帰係数を求めた (数式1)。回帰モデルについては印象トピックの影響度合いについての分析と同様に L1 ノルムを追加した正規化回帰モデルとした。

*1 一言コメント感情分析 API <http://nicomment.ktty1220.me/api.html>

label	ラベル	係数
abusive	暴言	0.0153265483
advice	助言	0.0440652424
critic	批判	-0.0024100129
cry	感動	-0.0032953198
exciting	期待	0.0609861821
fear	恐怖	0
feed	ツッコミ	-0.0023713537
fever	熱狂	-0.0003919834
greeting	挨拶	-0.0011993533
irritate	イライラ	-0.0020997182
like	賞賛	-0.0000224519
lol	笑い	0.0037913468
question	疑問	0.0018428950
request	要望	0.0003107976
response	返事	0.0147684423
surprise	驚愕	-0.0004399626

表3 視聴離脱回帰モデルの各感情の係数

結果

構築した回帰モデルの各感情係数は表3に示す。結果から、「なるほど」や「ですね」といったコメントが分類される「Response」, 「してください」や「してほしい」といったコメントが分類される「Request」, 「へたくそ」や「馬鹿」といったコメントが分類される「Abusive」が強く視聴離脱に影響していることが明らかになった。

結果の解釈

番組実況において、視聴者の意図する行動と演者の行動に差異が生じると、「～したほうがよい」のような「advice」を行うコメントが書き込まれる。「恋愛リアリティ」というジャンルでは、視聴者の多くが演者に対して感情移入しているため、「advice」のコメントが増える、つまり視聴者の意図と演者の行動が不一致になり、視聴離脱が発生すると考えられる。

7. おわりに

本研究では番組のシーンごとの印象や感情という「捉え方」の抽出および抽出した「捉え方」と視聴離脱との関連性について分析を行った。結果として、特定の印象と視聴離脱との間に相関があることが明らかになった。また、シーンに含まれる視聴者の投稿コメントが、ポジティブまたはネガティブが多数かつコメント数全体が多いもしくはポジティブとネガティブが拮抗している場合において、視聴離脱が特に発生するという傾向も明らかになった。そして、番組実況のようにインターネット用語や抽象的表現、タイピミスが含まれる自然言語においては、分かち書きを行い特定の品詞を抽出し、その内容について扱うよりも文字単位で全てのコメントを取り扱う方が良いと考えられる。

本研究では、「恋愛リアリティ」というコメントに視聴者の感情が多く含まれているジャンルについて分析を行っ

た。今後は、抽出しにくいジャンルについても拡張する必要がある。また、番組のシーンの連続性をふまえ、コメントの時系列を考慮することも重要であると考えられる。本研究では、番組のシーンと番組実況の抽出した「捉え方」を1対1で対応させていたが、今後は時系列生を考慮した実験と考察を試みる必要があるだろう。

参考文献

- [1] 映像コンテンツの視聴実態 - MarkeZine, 入手先 <<https://markezine.jp/article/detail/27922>> (2018.09.15).
- [2] 金多賢, 北島宗雄, 李登姫, 映像に対する嗜好と感情反応・印象評価の関係, 日本感性工学会論文誌, 2014.
- [3] ニールセン、生活者のマルチスクリーン利用動向を分析 テレビ視聴中にスマートフォンで行われているのはゲームやチャット入手先 <https://www.netratings.co.jp/news_release/2013/11/newsrelease20131112.html> (2013.11.12).
- [4] Blei, David M. and Ng, Andrew Y. and Jordan, Michael I., Latent Dirichlet Allocation, J. Mach. Learn. Res., 2003
- [5] Xiang Zhang and Junbo Jake Zhao and Yann LeCun, Character-level Convolutional Networks for Text Classification, CoRR,2015,abs/1509.01626
- [6] Robert Tibshirani, Regression Shrinkage and Selection Via the Lasso, Journal of the Royal Statistical Society, Series B,1994
- [7] 情報通信政策研究所, 総務省 「平成 29 年情報通信メディアの利用時間と情報行動に関する調査報告書」の公表入手先 <http://www.soumu.go.jp/menu_news/s-news/01iicp01_02000073.html>
- [8] 山内崇寛, 林佑樹, 中野有紀子, 日本語文解析による Twitter の感情分析とシーンインデキシングへの応用, 第 75 回全国大会講演論文集, 2013
- [9] 宮森 垣, 中村 聡史, 田中 克己, 番組実況チャットを利用したテレビ番組のメタデータ自動抽出方式, 情報処理学会論文誌,2005
- [10] Qing Ping, Chaomei Chen, Video Highlights Detection and Summarization with Lag-Calibration based on Concept-Emotion Mapping of Crowd-sourced Time-Sync Comments, 2017,1708.02210
- [11] 重森万紀, テレビとネット動画, 人々はどう使い分けているか テレビとネット動画, 人々はどう使い分けているか, 入手先 <<https://www.nhk.or.jp/str/publica/rd/rd158/PDF/P14-27.pdf>>