

動画に対する評価項目別スコアを用いた 直感的動画検索システム

嶋田 和士[†] 上田真由美^{††,†††} 中島 伸介[†]

[†] 京都産業大学 情報理工学部 〒603-8555 京都府京都市北区上賀茂本山

^{††} 流通科学大学 経済学部 〒651-2188 兵庫県神戸市西区学園西町 3-1

^{†††} 大阪大学 サイバーメディアセンター 〒560-0043 大阪府豊中市待兼山町 1 番 32 号

E-mail: [†]{g1853554,nakajima}@cc.kyoto-su.ac.jp, ^{††}Mayumi_Ueda@red.umds.ac.jp

あらまし 近年、動画共有サイトには膨大な動画が存在し、多くのユーザはタイトルやサムネイルから好みの動画を判断して視聴する。しかし、既存の情報では実際に再生するまでユーザが真に求めている動画か判断することが困難である。そこで、本研究では、動画に対して投稿されたコメントを分析することで、その動画に対するユーザの直感的評価を評価項目別スコアとして付与する。この評価項目別スコアを視聴前に提示することで、ユーザの直感に基づく新たな切り口による検索を可能とし、ユーザは動画を再生する前に、真に求めている動画か否かを判断可能となる直感的動画検索システムの開発を行なった。本稿では、提案手法の概要を述べるとともに、開発した直感的動画検索システムを用いた評価実験結果を報告する。

キーワード コメント分析, 直感的動画検索, 自動スコアリング, インターリーピング

1 はじめに

近年、動画共有サイトには膨大な動画が存在し、多くのユーザはキーワード検索をした後に、タイトルやサムネイルから好みの動画を判断して視聴する。しかしながら、図 1 に示す通り、ユーザが直感的に観たいと思える直感的動画検索は、従来のキーワード検索では限界がある。例えば、エモい動画が観たくてエモいとキーワード検索をしても、タイトルや説明文にエモいというワードが含む動画しかヒットしない。実際にエモいと感じる動画は、タイトルにエモいと入ってなくても、ユーザのコメントとしてエモいと同様な言葉である、懐かしいや切ないといったコメントが見られる。そのため、直感的動画検索は従来の単純なキーワード検索では限界がある。また、既存の情報では実際に再生するまでユーザが真に求めている動画か判断することが困難である。

そこで本研究では、ユーザが直感的に観たいと思える直感的動画検索を可能とするシステムの開発を目指し、まずは評価項目別コメント自動スコアリング手法の提案を行う。本手法では動画に対して投稿されたコメントを分析することで、その動画に対するユーザの直感的評価を評価項目別スコアとして付与する。そして、この評価項目別スコアを視聴前に提示することで、ユーザの直感的に基づく新たな切り口による検索を可能とし、ユーザは動画を再生する前に、真に求めている動画か否かを判断可能となる手法を提案する。本稿では、提案手法のシステムについて説明すると共に、提案システムの有用性を検証した評価実験について報告する。



図 1 従来の動画検索手法の問題点

2 関連研究

近年の動画共有サイトにはユーザが自由にコメントでき、YouTube [1] やニコニコ動画 [2] などが有名である。また、上記サイトには膨大な動画が日々アップロードされている。ユーザはタグ検索やキーワード検索で観たい動画を探し、動画に対する評価・コメントを自由に書き込むことができる。

動画分析に基づく関連研究として、村上らは動画に付与されるコメントの文字「w」に着目し、解析を行うことで、動画ごとの「笑い」の度合いを数値化する手法を提案した [3]。中村らは、感性に基づく正規表現辞書を構築し [4]、ユーザが書き込んだコメントを分析して、positive, negative, happy, sad, surprise といった感性に基づくランキング検索を可能とした [5]。本研究では、これらの感性分析には留まらず、ユーザのより直感的な動画検索を可能とする技術の開発を目指している。

また、レビュー分析に基づく評価表現別自動スコアリング手法に関する従来研究として、松波らは、コスメアイテムに特化した評価表現辞書を構築し、レビュー自動スコアリング手法を開発した [6]。同様に、安部らは、ホテル予約サイトに投稿さ

れたレビュー文を分析することで、ホテルに特化した評価表現辞書を構築し、ホテルレビュー自動スコアリング手法を開発した[7]。本研究ではこれらの辞書構築手法を参考にしつつも、動画に対するより直感的な評価項目の選定と各評価項目に対する評価表現の効率的な抽出を目指すものである。

上述の通り、動画のコメント分析による感性検索の研究や、レビュー解析による評価表現別のスコアリング手法に関する研究は数多く行われているが、本研究にて実施している直感的動画検索のための評価表現別自動スコアリング手法に関する研究開発は行われていない。

3 評価項目別コメント自動スコアリング方式

3.1 提案手法の概要

我々は、動画に付与されるコメントを分析し、評価項目別自動スコアリングをすることで、動画を再生する前に、真に求めている動画か否かを判断可能となる手法を提案する。図2のように、従来手法では動画共有サイトの一覧から主にタイトルとサムネイルを見て動画を選択するため、予想していた内容と異なることがある。一方、提案手法では、各動画に、従来のキーワード検索では拾えないような、また、ポジティブやネガティブと言った単純な感性検索ではない動画に特化した評価項目にスコアを付与することで、視聴前からユーザが真に求めている動画か否かを判断可能とする。



図2 提案手法の概要

3.2 評価項目の検討

本節では、動画に対する評価項目を検討する。本研究では、楽しいや悲しいと言った単純な感性検索ではなく、動画に特化した直感的な検索を可能とすることを目的としている。本稿で検討する評価項目は、エモい、チル、きゅん、じわる、ぴえん、腹立ち、グロい、映え、萌えの9つとする。今後も評価項目は増やしていく予定である。

3.3 各評価項目の分類器モデルの構築

本節では、動画に対する評価項目別のスコアを有する評価項目別分類器モデルの構築について述べる。動画を視聴するユーザは多種多様な直感的評価表現で動画に対するコメントを投稿している。そこで、実際に使用されている直感的評価表現を収集し、学習データに使用することで評価項目別分類器モデルの構築を行う。

学習データに使用するコメントデータは、動画共有サイトYouTubeが提供するYouTube Data API [8] から取得する。分類器モデルの構築にはfastText [9] を用いる。例としてエモい分類器モデル構築についての手順を示す。

エモい分類器構築手順

- 1 評価項目だと感じる動画を収集し、各評価項目ごとに5000コメントずつ収集する。
- 2 fastTextを用いて評価項目9つに対してマルチラベル分類モデルを作成する。
- 3 エモくないコメントを収集するため、2で構築した分類器を用いて200本の動画をスコアリングする。
- 4 エモいスコアが0.05以下の動画をエモくない動画と定め、そこから5000コメントを収集する。
- 5 エモいコメント5000とエモくないコメント5000を用いてfastTextで学習させエモい分類器モデルを構築する。

上記手順同様に、各評価項目ごとに9つの分類器モデルを構築する。図3のように、未知なる動画のコメントを一まとめにして、各分類器に通すと評価項目ごとのスコアが算出される。

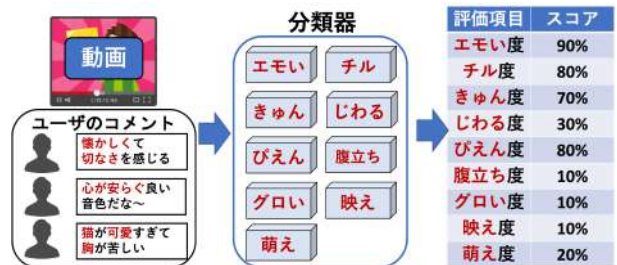


図3 提案手法の手順

4 直感的動画検索システムの開発

本章では、3章での提案手法を基に開発した、直感的動画検索システムについて述べる。

4.1 直感的動画検索システムで使用する動画データ

本システムにおいて格納している動画本数は、700本である。詳細な動画の内訳は下記である。

- ・ YouTubeで評価項目名（例：エモい）と検索し、関連度順で上から10本の動画 × 評価項目数9により、計90本
- ・ 手で評価項目名だと感じる動画を610本
- ・ コメント数が30以上ある動画

4.2 直感的動画検索システムの使い方

図4のように、「検索する評価項目を選択」から検索したい評価項目を選択し、「検索する評価項目の優先度」に移動させ、検索することで図5の検索結果画面になる。図5はエモいスコアが高い順に表示され、各スコアとレーダーチャートが付与される。これらがあることで、各動画同士を比較しやすくし、動画を視聴する前からどのような動画が分かるようになる。また、検索手法は評価項目が2つ以上や、キーワードと評価項目を含めた検索もできる。



図 4 直感的動画検索システムの検索画面



図 5 直感的動画検索システムの検索結果画面

に、生成したエモい混合ランキング TOP10 を被験者に提示し、直感的にエモいと感じる動画をクリックし、視聴してもらう。この時ユーザに提示する情報はサムネイルとタイトルのみである。また、視聴する際の注意点として、基本的に動画は最後まで観るようにお願いした。ただし、エモいと感じた時点で視聴はやめても構わないことにし、作業用 BGM 等の長い動画は 5 分～10 分程度飛ばし飛ばしで見てもらい良いことにした。図 8 のように、視聴後にエモいと感じたかを 5 段階評価で送信してもらう。「はい」または「ややはい」の場合、credit を与えることにする。これらを直感的にエモいと感じる動画がなくなるまで繰り返してもらう。以上、同様にチル、きゅん、じわる、びえん、腹立ち、映え、萌えの評価項目ごとに行ってもらい。また、今回評価項目グロいは倫理問題上、実験を行わないことにした。



図 6 インターリービング実験の概要

5 評価実験

本章では、3.3 節にて述べた自動スコアリング方式による従来システムとのランキング性能の比較評価実験 (5.1) と 4 節で構築した直感的動画検索システム自体の評価実験計画 (5.2 節) について述べる。

5.1 ランキング性能比較実験

本節では、従来手法のランキング性能と提案手法のランキング性能の比較実験について述べる。

5.1.1 実験参加者

被験者は男性 18 人、女性 1 人の計 19 人の被験者に対して実験を行った。また、被験者の年代は 19 人全員が 20 代であった。

5.1.2 実験手順

従来手法 (YouTube) と提案手法 (直感的動画検索システム) でのランキング性能を比べるため、インターリービング手法を用いる。インターリービングを用いた評価手法は飯塚洗二郎ら [10] や、加藤誠ら [11] の研究を参考にした。例として、エモいランキングの評価方法について述べる。図 6 のように、従来手法で「エモい」とキーワード検索をし、関連度順に上から 10 本の動画を従来手法のエモいランキング TOP10 とする。また、提案手法でエモい検索をし、エモいスコアが高い順に上から 10 本の動画を提案手法のエモいランキング TOP10 とする。この時の動画セットは 4.1 節の通りである。この 2 つのランキングからエモい混合ランキングを生成する。混合ランキング生成手法は Probabilistic interleaving [12] を用いる。図 7 のよう



図 7 エモい混合ランキングの提示



図 8 視聴後の評価項目判別送信画面

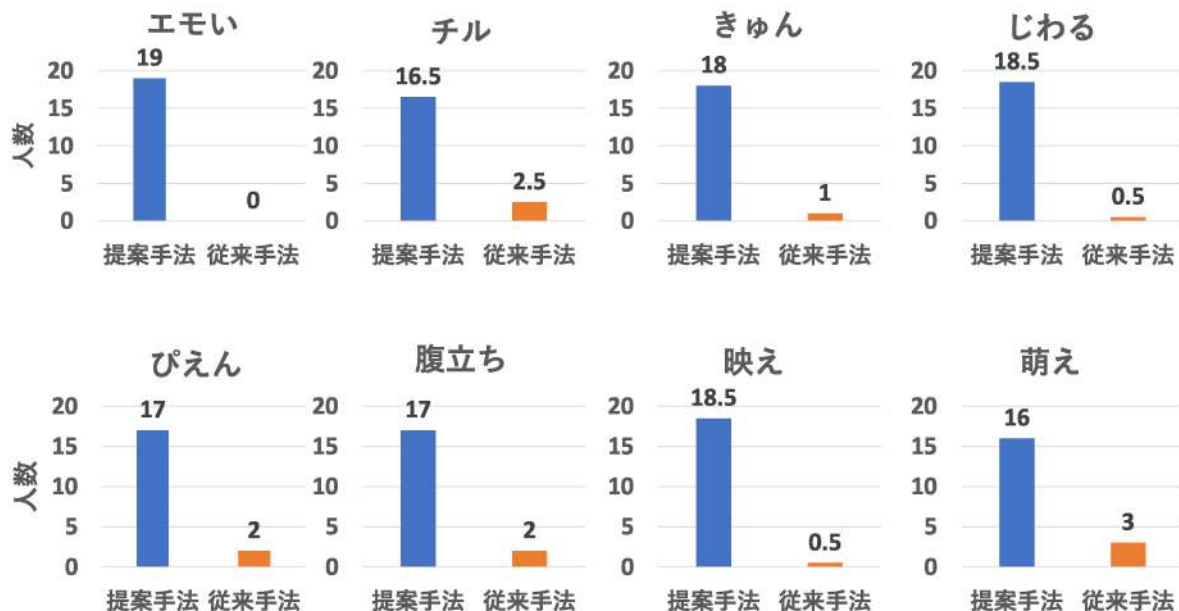


図 9 評価項目ごとの Probabilistic interleaving 結果

	エモい	チル	きゅん	じわる	ぴえん	腹立ち	映え	萌え
p値	4.33×10^{-28}	1.45×10^{-3}	6.64×10^{-7}	5.89×10^{-13}	4.74×10^{-6}	4.74×10^{-6}	5.89×10^{-13}	7.15×10^{-5}

図 10 評価項目ごとの Probabilistic interleaving の p 値

5.1.3 実験結果

図 9 は評価項目ごとの Probabilistic interleaving の結果である。引き分けの場合、従来手法・提案手法どちらにも 0.5 人数付与する。エモいランキング結果は、従来手法が 0 人、提案手法が 19 人となり、提案手法のエモいランキングが優れているという結果となった。以下同様に、チルランキング結果は、従来手法が 2.5 人、提案手法が 16.5 人となり、提案手法のチルランキングが優れているという結果になった。きゅんランキング結果は、従来手法が 1 人、提案手法が 18 人となり、提案手法のきゅんランキングが優れているという結果になった。じわるランキング結果は、従来手法が 0.5 人、提案手法が 18.5 人となり、提案手法のじわるランキングが優れているという結果になった。ぴえんランキング結果は、従来手法が 2 人、提案手法が 17 人となり、提案手法のぴえんランキングが優れているという結果になった。腹立ちランキング結果は、従来手法が 2 人、提案手法が 17 人となり、提案手法の腹立ちランキングが優れているという結果になった。映えランキング結果は、従来手法が 0.5 人、提案手法が 18.5 人となり、提案手法の映えランキングが優れているという結果になった。萌えランキング結果は、従来手法が 3 人、提案手法が 16 人となり、提案手法の萌えランキングが優れているという結果になった。よって、全ての評価項目が提案手法の方が優れているという結果になった。また、図 10 に t 検定 (対応あり, 両側, 有意水準 5%) の各評価項目ごとの p 値を示す。このことから、全ての評価項目において有意差があることが確認できた。

5.1.4 考察

Probabilistic interleaving の結果により、全ての評価項目において提案手法のランキング性能の方が従来手法より優れているという結果になった。また、全ての評価項目において有意差も見られた。このことから、従来手法で単純に評価項目名で検索して出てくる動画以上に評価項目名だと感じる動画は他にあるということが判明した。単純にタイトルや説明文に評価項目名が書いてある動画以上に、評価項目名と感じる直感的評価表現のコメントがある動画の方が、ユーザが真に求める動画であるということが確認できた。

5.2 直感的動画検索システム評価実験

本節では、直感的動画検索システム自体の評価実験について述べる。

5.2.1 実験参加者

被験者は男性 14 人の被験者に対して実験を行う。また、被験者の年代は 14 人全員が 20 代である。

5.2.2 実験手順

例として、エモい検索のシステム評価実験方法について述べる。

- 1 従来手法 (YouTube) で「エモい」とキーワード検索をしてもらい、直感的に見たいと思う動画を 1 つ見てもらう。
- 2 4 節で述べた提案手法 (直感的動画検索システム) を用いて、「エモい」検索をしてもらい、サムネイル・タイトル・スコア・レーダーチャートを参考に直感的に見たいと思う

動画を1つ見てもらう。

- 3 図11に示す独自アンケートに回答してもらう。従来システムとの比較評価になっており、回答は「提案システムの方が良い」「どちらかというとな提案システムの方が良い」「従来システムと変わらない」「どちらかというとな従来システムの方が良い」「従来システムの方が良い」の5つから選択できる。順に5点、4点、3点、2点、1点とする。

A	観たい動画を探しやすい
B	視聴前から動画の評価を比較しやすい
C	視聴前と視聴後で動画に対する印象のギャップがない
D	各動画に対する評価がわかりやすい
E	気になる動画が多くあった

図11 評価実験のアンケート

上記手順を各評価項目ごとにやってもらう。この時、1 従来手法→2 提案手法のグループと2 提案手法→1 従来手法のグループに分けて行う。全評価項目について独自アンケート終了後、従来手法・提案手法それぞれについて System Usability Scale [13] を行ってもらう。質問項目は図12である。最後に自由記述で直感的動画検索システムについて良かった点と改善点を回答してもらう。

1	このシステムは直感的検索ができる
2	このシステムは不必要に複雑だった
3	このシステムは視聴前と視聴後で動画に対する印象のギャップがない
4	このシステムを利用するには、技術者のサポートが必要だと思う
5	このシステムを頻繁に使用したいと思う
6	このシステムは使うのがとても面倒だと感じる
7	このシステムは使いやすいと感じた
8	このシステムでは一貫性のないところが多々あったと感じた
9	自信を持ってこのシステムを操作できた
10	このシステムを使いこなすには事前にたくさんの知識が必要だと思う

図12 System Usability Scale の質問項目

5.2.3 実験結果と考察

独自アンケートによる従来手法との比較結果を図13に、被験者ごとの SUS による従来手法との比較結果を図14に示す。

まず、独自アンケートによる従来手法との比較結果を分析する。図13より、全体的に提案システムの方が良いと答えた人が多く、従来システムよりも直感的検索ができたことが分かる。

また、図11の3「視聴前と視聴後で動画に対する印象のギャップがない」について、エモイ・チル・映え・萌えの平均点は3点台なので従来システムと変わらないという結果になり、サムネイルとタイトル通りの動画が多いことが分かった。反対に、きゅん・じわる・ぴえん・腹立ちの平均点は4点台、または4点近いので提案システムの方が良い結果となり、サムネイルとタイトルだけでは視聴前と視聴後で動画に対する印象にギャップが生まれる動画が多いことが分かったのと同時に、評価項目ご

とにスコアを提示することによりギャップを無くすることができた。評価項目ごとにギャップが生じるかどうかの原因は、エモイ・チルは曲系統、映えはオシャレなカフェ紹介動画系統、萌えは女性キャラのアニメ系統で一貫してサムネイル通りのイメージの動画が多く、動画の途中で印象が変わるものは少なかったと考える。また、きゅん・じわる・ぴえん・腹立ちはサムネイルとタイトルの印象とは別に、動画の途中で印象が変化する動画も数多く含まれていたと考える。

次に、SUS 結果についてを分析する。図14より、3人のみ従来手法の方が得点が高くなったが、SUS における従来手法の平均点は73.58、提案手法の平均点は79.29となった。よって、従来手法よりユーザビリティが優れていることが分かった。ただし、t検定(対応あり、両側、有意水準5%)において有意差は見られなかった。また、図14の3「このシステムは視聴前と視聴後で動画に対する印象のギャップがない」に着目すると従来手法の平均点が3.14、提案手法の平均点が4.21であった。また、t検定(対応あり、両側、有意水準5%)において有意差が見られた。このことから、全体を通してこの直感的動画検索システムは従来システムより視聴前からユーザが真に求めている動画か否かを判断可能であることが分かった。最後に、自由記述による提案システムの良い点と改善点を以下にまとめる。

良い点

- ・レーダーチャートにより動画を比較しやすく、探しやすい
- ・実際に動画を見た際にもギャップが少なかった
- ・評価項目名と感情を抱くことが理解できる動画が従来手法より多かった
- ・従来手法より検索結果の偏りが少なかった

改善点

- ・インターフェースの改善
- ・ユーザの要望に合った柔軟な検索方法の実現
- ・各評価項目に対するユーザの印象の違いを考慮

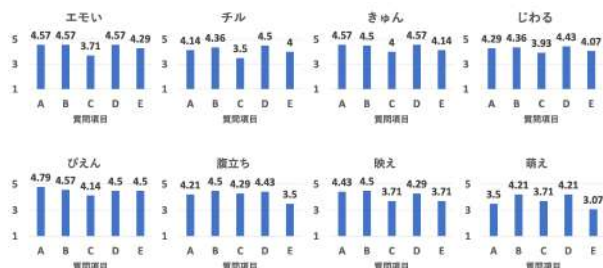


図13 独自アンケートの結果

6 おわりに

本稿では、ユーザの直感に基づく新たな切り口による検索を可能とし、動画を再生する前に、真に求めている動画か否かを判断可能となるシステムの開発を目的とした。その実現の為に直感的動画検索システムを提案した。評価実験の結果、提案した評価項目全てにおいてランキング性能が従来手法より優れていることが確認でき、スコアを提示することによって視聴前と

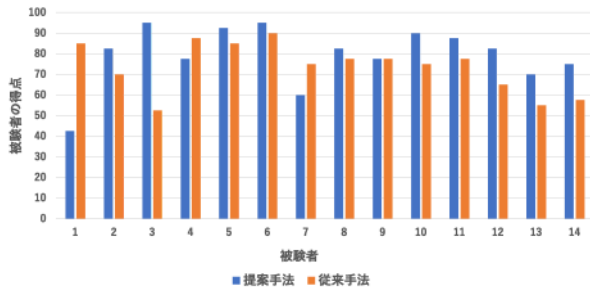


図 14 被験者ごとの SUS 結果

視聴後のギャップをなくすことが確認できた。

今後は、新たな評価項目の追加や、自動スコアリングの精度向上にも努めていきたい。

謝 辞

本研究の一部は、JSPS 科研費（20H04293, 19K12243）および京都産業大学先端科学技術研究所（ヒューマン・マシン・データ共生科学研究センター）共同研究プロジェクト（M2001）の助成を受けたものである。ここに記して謝意を表す。

文 献

- [1] YouTube <https://www.youtube.com/>
- [2] ニコニコ動画 <https://www.nicovideo.jp/>
- [3] 村上 直至ら，動画コンテンツの視聴者コメントに基づくランキングとその評価，DEIM Forum 2012.
- [4] Hisashi Miyamori, et al., Generation of Views of TV Content Using TV Viewers' Perspectives Expressed in Live Chats on the Web, ACM Multimedia 2005, 2005.
- [5] 中村 聡史ら，ソーシャルノテーションに基づく動画検索手法，DEIM Forum 2009 D6-1, 2009.
- [6] 松波 友稀ら，コスメアイテムに対する評価項目別レビュー自動スコアリング方式の開発，DEIM Forum 2017 B5-3, 2017.
- [7] 安部 克ら，レビュー評価項目別スコアを用いたホテル推薦手法の提案，DEIM Forum 2019 P1-134, 2019.
- [8] YouTube Data API
<https://developers.google.com/youtube/v3/getting-started>
- [9] fastText <https://fasttext.cc/>
- [10] 飯塚 洸二郎ら，ニュースアプリケーションのパーソナライゼーションアルゴリズムに対するマルチリーピング手法の比較，The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019 2L3-J-9-05, 2019.
- [11] 加藤 誠ら，コミュニティQ&A サイトにおける質問検索システムの大規模オンライン評価，DEIM Forum 2018 G2-2, 2018.
- [12] Katja Hofmann, Shimon Whiteson and Maarten de Rijke, A Probabilistic Method for Inferring Preferences from Clicks, Proceedings of the 20th ACM Conference on Information and Knowledge Management, CIKM 2011, Glasgow, United Kingdom, October 24-28, 2011.
- [13] Brooke, J. 1996. "SUS: A "quick and dirty" usability scale". In Usability evaluation in industry, Edited by: Jordan, P. W., Thomas, B. A. Weerdmeester and McClelland, I. L. 189-194. London: Taylor & Francis.