SmartVideoRanking: 視聴者の時刻同期コメントに基づく 動画ランキングシステム

佃 洸摂^{1,a)} 濱崎 雅弘^{1,b)} 後藤 真孝^{1,c)}

概要:本稿では、視聴者の反応をクエリとして使用することで、ユーザが自分好みの動画のランキングを生成できるシステム SmartVideoRanking を提案する。たとえばユーザが「初音ミク」というクエリを入力すると、SmartVideoRanking は「ベースかっこいい」や「鳥肌がすごい」といった視聴者の反応をキーワードとしてユーザに推薦する。ユーザが「鳥肌がすごい」というキーワードを選択すると、SmartVideoRankingは「初音ミク」に関する動画の中で「鳥肌がすごい」という反応が多い動画ほど上位に表示されるランキングを生成してユーザに提示する。本研究では、クエリに対して有用な視聴者の反応を抽出するために、視聴者が動画に投稿した時刻同期コメントを利用する。提案手法では、コメントに関する特徴を用いた機械学習により、クエリに対するコメントの有用度を推定する。50個のクエリを用いた評価実験の結果、提案手法によって推定された有用度と評価者によって判定された有用度の相関は平均で 0.7547 を達成した。また、クエリに関する動画を再生数の多い順に表示する従来のランキングと、ユーザが選択したコメントに基づいて生成されるランキングを比較したところ、上位 30 件の重複の平均は 2.923 件であり、コメントを検索に用いることで従来のランキング方法では発見が困難な動画を閲覧可能なランキングを生成できることを示した。最後に、SmartVideoRanking のプロトタイプを実装し、コメントに基づく動画のランキングシステムを実現した。

1. はじめに

動画共有サービスや録画用デバイス,動画編集ツールの 普及に伴い、専門的な知識を持たない人々であっても動画 を作成し Web 上で発信することが容易になっている. これ により、Web 上で視聴できる動画の数が増加しただけでな く, 多様な動画が視聴できるようになってきた. 動画共有 サービス上で動画を視聴するユーザに目を向けると, ユー ザが視聴する動画を決めるための代表的な手段として検索 があげられる.動画検索では、ユーザの望む動画のタイト ルやタグに含まれる語をクエリとして入力し、検索された 動画集合を再生数やお気に入り数などの指標に基づいてラ ンキングすることが一般的である. しかし, 多くのユーザ はランキングの上位のみを閲覧するため [1], 限られたラン キング指標の中でユーザの目に触れるのはごく一部の動画 にとどまる. さらに, 動画コンテンツの場合, 文書に比べ てテキスト情報が少ないため,タイトルやタグに含まれる 語だけではユーザが望む動画の検索が難しい場合がある. これらの結果として, 多様な動画が投稿されているにも関

わらず,一部の動画のみが集中的に視聴され,大部分の動画はほとんど視聴されないという現象が起きている [2].

そこで本研究では、従来の指標に基づく動画のランキン グとは異なる,新たな動画ランキング方法およびその実現 のための手法を提案する.書籍や映画,家電といった製品 に対するレビューサイトが人気を集めていることや、レ ビュー情報を集約して提示することでユーザの意思決定の 支援を行うこと目的とした研究 [3], [4] が多数行われてき たことから,人々が製品を閲覧したり購入したりする際に, その製品に対して他の人がどのように反応しているかを 提示することは有用であると言える. 我々はこの考えを動 画検索に適用し、動画への視聴者の反応に基づく動画のラ ンキングの生成を目的とする. 具体的には, たとえば「初 音ミク」というクエリに対して,「ベースかっこいい」や 「鳥肌がすごい」といった反応をユーザに提示することで、 ユーザが「『ベースかっこいい』という反応の多さに基づく 『初音ミク』の動画ランキング」や「『鳥肌がすごい』という 反応の多さに基づく『初音ミク』の動画ランキング」を生 成することを可能にする. 動画への視聴者の反応を取得す るために、本研究では視聴者によって動画の再生時刻と同 期して投稿されたコメント (時刻同期コメント [5]) を利用 するというアプローチをとる. 我々は、時刻同期コメント

¹ 国立研究開発法人 産業技術総合研究所

a) k.tsukuda@aist.go.jp

b) masahiro.hamasaki@aist.go.jp

c) m.goto@aist.go.jp





コメントの有用度推定			
順位 コメント			
1	めっちゃかっこいい		
2	ベースかっけえ		
3	良い曲と思う		
4	調教うめえ		
:	:		

図1 本研究の概要.

にはランキングの生成に有用なコメントが記述されやすい と仮定し,動画共有サービスのニコニコ動画 *1 の時刻同期 コメントからクエリにとって有用なコメントを推定する.

動画共有サービスに投稿される動画のジャンルは様々であるが、代表的なジャンルのひとつに音楽があげられる。音楽動画の中には、ユーザ自らが作詞や作曲をした動画だけでなく、他のユーザが作った音楽を歌ったり、演奏したり、音楽に合わせて踊ったりする派生動画もある [6]、[7]. さらに、派生動画同士を合成することで新たな動画を作成する「N次創作 [8]」も盛んに行われている。N次創作動画では、その性質上タイトルの類似した動画が多くなりがちであり、上述の動画検索時の問題はより深刻になる。そこで、本稿では音楽動画とその派生動画を対象として評価実験およびシステムの実装を行う。

本研究の概要を図1に示す.本研究の主な貢献は次の4点である.(1)動画ランキングにおいて,視聴者のコメントを指標とするランキングという概念を提案した.(2)サポートベクターマシン(SVM)の線形回帰によりコメントの有用度推定を行った.50個のクエリを用いた実験の結果,推定された有用度と評価者が判定した有用度の相関は平均で0.7547を達成した.(3)再生数やお気に入り数などの従来のランキングでは上位に出現しない動画であっても,視聴者の反応に基づくランキングでは上位に出現し,ユーザがより多様な動画を閲覧できることを示した.(4)視聴者の反応に基づく動画ランキングシステムSmartVideoRankingを実装した.

本稿の構成は以下の通りである。2章では関連研究について述べる。3章では本研究が対象とする動画共有サービスであるニコニコ動画の概要および、本研究で使用するデータセットについて述べる。4章ではコメントの有用度の推定に用いる特徴を述べ、5章で評価実験について述べる。6章で視聴者の反応に基づく動画ランキングシステムSmartVideoRankingについて述べ、7章で本稿をまとめる。

2. 関連研究

2.1 コメントを用いた動画の分析・検索

動画の再生時刻とは非同期のコメントを用いた研究として、YouTube *2 の動画に投稿されたコメントを検索に利用した研究があげられる。Yee ら [9] および Chelaru ら [10] は、動画のタイトルとタグのみからインデックスを作成するよりも、タイトル、タグ、コメントを用いてインデックスを作成した方が検索精度が上昇することを示した。これらの研究では、クエリとの関連度に基づく従来のランキングをより高い精度で生成するためにコメントを利用しているのに対して、我々はクエリに関する有用な視聴者の反応の発見および視聴者の反応に基づく動画ランキングの生成のためにコメントを用いるという違いがある。

本研究に最も近い研究として, ニコニコ動画のコメントを 活用した動画検索の研究があげられる. Nakamura ら [11] は, positive, negative, happy, sad の4種類の印象の強さ に基づいて動画をランキングする手法を提案した. 彼らの 手法では、各印象を表す際に記述される語の集合を辞書と して事前に人手で作成し,動画中のコメントと辞書内の語 のマッチングを行うことで、各動画の各印象の強さを求め ている. 佃ら [12] は、Nakamura ら [11] の作成した辞書の 使用に加えて, コメント中のキャラクタ名を抽出すること で、「初音ミクというキャラクタが動画の後半に活躍する 泣ける動画」のような検索を実現した. Yamamoto ら [13] は、時刻同期コメントの特徴を用いた機械学習により、音 楽動画を「可愛らしい」、「切ない」、「元気がでる」といった 印象に分類する手法を提案した. また実験により, 楽曲の 歌詞や音響特徴量を用いた手法よりも時刻同期コメントを 用いた手法の方が高い精度で印象の分類ができることを示 した. これらの研究では、動画を視聴したユーザがどう感 じたかという情報に基づいた動画の検索の実現を目的とし ているが、検索に使用できる印象は事前に用意されたもの に限られている. それに対して本研究では, クエリに応じ た,有用な視聴者の反応を時刻同期コメントから発見する.

ニコニコ動画は日本に固有のサービスであるが、海外であっても、テレビ番組を視聴しながら Twitter *3 にコメントを投稿することは一般的に行われている。そのようなコメントをテレビ番組に対する時刻同期コメントとみなし、動画の盛り上がりや、ポジティブ・ネガティブといった動画に対する視聴者の印象を分析する研究 [14]、[15] が行われてきた。これらの研究では動画の分析を目的としているが、動画コンテンツが今後ますます増加していく中で、時刻同期コメントを検索に活用することは、ユーザが動画を視聴するうえで大きな役割を果たすと我々は考える。

^{*1} http://www.nicovideo.jp

^{*2} http://www.youtube.com

^{*3} http://twitter.com

2.2 クエリ推薦

クエリにとって有用な視聴者の反応をユーザに提示する 本研究は、クエリ推薦に関する研究であるとも言える. ク エリを推薦する際の主なアプローチには、検索精度を高め るために元のクエリを置換する方法 [16], [17] (たとえば 「vehicle wash」というクエリに対して「car wash」を推薦) と検索の対象を明確にするために元のクエリにキーワード を追加する方法 [18], [19] (たとえば「michael jordan」とい うクエリに対して「michael jordan shoes」や「pictures of michael jordan」を推薦)がある.後者のような推薦を実 現するために、検索ログ内でクエリおよび各クエリの検索 結果中でクリックされた URL を基に二部グラフを作成し, 二部グラフ上で値を伝播させることでクエリの関連度を 求める手法が多数提案されてきた. 本研究は後者との関連 が強いが、これまでに提案された手法を用いるのは困難で あると考えられる. その理由として, 現在の動画検索サー ビスは視聴者の反応に基づく検索をサポートしておらず, 我々が発見したいような視聴者の反応に関するキーワード は一般に入力されないため, クエリログを用いてもスパー スな二部グラフしか作成されないことがあげられる. この ため, 我々は動画に投稿された時刻同期コメントから有用 な視聴者の反応を発見する手法を提案する.

3. ニコニコ動画とデータセット

3.1 ニコニコ動画

ニコニコ動画は、2015年8月の時点で登録会員数5,000 万人,投稿動画数 1,200 万件を超える,日本で最も利用さ れている動画共有サービスのひとつである. ニコニコ動画 の最も特徴的な機能のひとつはコメント機能にあり、ユー ザは動画の再生時刻に同期したコメント(時刻同期コメン ト[5])を投稿できる.動画の下にリスト形式でコメントが 表示される一般的な動画共有サービスとは異なり、ニコニ コ動画ではコメントが投稿された再生時刻になると,他の ユーザが投稿したコメントと共に画面上を右から左に移動 しながら表示される. これによりユーザは, 実際には別々 の時間に動画を視聴していても, あたかも他のユーザと同 時に動画を視聴しているかのような体験ができる [8]. 動 画の再生時刻に非同期のコメントに比べて, 時刻同期コメ ントは「その瞬間ごとの感情をより多く表現する傾向があ る [5]」ことが指摘されている. そのため, たとえば多くの 人が面白いと感じた時刻付近では「面白い」や「楽しい」 などのコメントが大量に表示される. 本研究では時刻同期 コメントが持つこの性質に着目し,動画のランキング生成 のための視聴者の反応を時刻同期コメントから抽出すると いうアプローチをとる.

3.2 データセット

本研究では,国立情報学研究所により提供されている

「ニコニコデータセット *4」を用いる。このデータセットには、2012 年 11 月初旬までにニコニコ動画に投稿された約 830 万件の動画のメタデータおよび、各動画に投稿されたコメントデータが含まれる。動画のメタデータの中で、本研究で使用するのはタグ、再生数、コメント数、マイリスト数である。コメントデータについては、本研究ではコメントの本文、コメントが投稿された日時、コメントが投稿された動画上の再生時刻の情報を使用する *5.

約830万件の動画の中から、本研究では VOCALOID のオリジナル楽曲動画およびその派生動画を対象とする.派生動画は、歌ってみた、踊ってみた、演奏してみた、PVをつけてみた、MMD *6をつけてみた、の5カテゴリのいずれかに含まれる動画を対象とした。オリジナル動画および派生動画のカテゴリの判定には、Hamasakiら [20] が公開している音楽視聴支援サービス Songrium *7の分析データを用いた。さらに、本研究ではコメントが 1,000 件以上投稿されている動画のみを対象とし、コメントデータは投稿日時の新しい 1,000 件を対象とした。最終的に作成されたデータセットに含まれる動画数は 11,180 件、コメント数は 1,118 万件であった。

4. 手法

本研究では、検索クエリに対するコメントのランキング 指標としての有用度を SVM の線形回帰により求める.線 形回帰を用いたのは、学習データの数に対して特徴の数が 大きいためである.以下、4.1節ではコメントの特徴を求め るうえで必要な前処理について述べる.4.2節では有用度 を推定する対象となるコメントの収集方法について述べ、 4.3節では有用度の推定に使用する特徴について述べる.

4.1 コメントの前処理

4.1.1 正規化

Brody ら [14] は、Twitter に投稿されるコメントには、コメントが持つ意味を強調するために「coooooollllll」のように単語内の文字が繰り返して記述される傾向があることを報告している。ニコニコ動画のコメントでも同様の傾向が見られ、「かわいい」という気持ちをより強く表現するために、「かわいいいいいい!!!」のように文字や記号を繰り返したコメントが存在する。つまり、元々は同じ意味を表すコメントであるにも関わらず、繰り返されている文字の数が異なるために別の文字列となるようなコメントが多数存在する。コメントの有用度を求める際に、そのようなコメントを別々に扱うと、データセット内での各コメントの出現数が少なく、特徴量を計算するうえで不十分であることが起こりうる。Brodyら [14] は、このような表記

^{*4} http://www.nii.ac.jp/dsc/idr/nico/nico.html

^{*5} コメントを投稿したユーザ ID はデータ提供時に削除されている.

^{*6} http://www.geocities.jp/higuchuu4/

^{*7} http://songrium.jp

揺れの問題を解決するために、コメントの正規化手法を提 案した. Yamamoto ら [13] も, ニコニコ動画のコメントを 用いて動画の印象推定を行う際に、Brody ら [14] の手法に 倣ってコメントを正規化した. 本研究でもこれらの研究に 倣い,以下の手順でコメントの正規化を行う.(1)スペー スおよび、日本語と英数字以外の記号(「!」や「+」など) を除去する. (2) 半角の片仮名および英数字を全角に変換 する.(3)平仮名および片仮名の小書き文字(「ぁ」「ァ」 「っ」「ッ」など)を大文字に変換する.(4)英語の小文字 を大文字に変換する. (5) ニコニコ動画のコメントでは, 楽しい気持ちを表す際に「w」が用いられることが多い. 「すごい」と「すごいww」を同じ反応として扱うために、 コメントの末尾の「w」を全て除去する. (6) 2回以上 繰り返して記述されている文字の繰り返し回数を1回にす る. たとえば「かわいい」や「かわいいいいいい」は「か わい」に変換される.以上の(1)から(6)の操作によ り、たとえば「かわいい」と「かわぃぃいいいいい!!!」 はともに「かわい」という同じ文字列に正規化される.

4.1.2 正規化コメントの復元

「かわいい」というコメントを正規化すると「かわい」になるように、正規化コメントは日本語として不自然になることも多く、ランキングの指標としてユーザに提示するのは不適切である.そこで、正規化してcという文字列になるコメントの中で、データセット中で最も出現頻度の高いコメントをcの典型的な元コメントとして c_t で表す.正規化コメントcがランキング指標として有用であると推定されユーザに提示する際は c_t を提示する.

4.2 コメントの収集

クエリを q としたとき,タグに q を持つ動画を q と関連 のある動画とみなし,そのような動画の集合を V_q とする. 4.1 節で述べたコメントの正規化を行ってもなお出現数の 少ないコメントを除くために, V_q 内の 3 件以上の動画に投稿されており,かつ総出現数が 10 回以上の正規化コメントを有用度の推定対象とし, C_q で表す.

4.3 特徴

本節では、コメント c_t (正規化コメント c)に対して本研究で用いる特徴について述べる。以下の特徴の中で、使用するのが音楽動画のコメントに限られるのは 4.3.9 項および 4.3.13 項の特徴のみであり、その他の特徴は任意の動画のコメントに対して使用可能である。

4.3.1 クエリ関連度(QREL)

クエリとの関連度が高いコメントの有用度は高いと仮定し、次式で計算される Expected Mutual Information [21] の値を特徴として用いる.

$$epmi(q, c_t) = P(q, c) \cdot log \frac{P(q, c)}{P(q)P(c)}.$$
 (1)

ここで,
$$P(q)=rac{|V_q|}{|V|}$$
, $P(c)=rac{|V_c|}{|V|}$, $P(q,c)=rac{|V_q\cap V_c|}{|V|}$ であ

り、V はデータセット内の全動画集合、 V_c は V の中で c が 1 件以上投稿されている動画集合を表す.

4.3.2 クエリ類似度(QSIM)

「初音ミク」というクエリに対する「ミクかわいい」のように、クエリと文字の重複があり、クエリとの類似度の高いコメントの有用度は高いと仮定し、編集距離の一種である正規化レーベンシュタイン距離 [22] を用いて次式で計算される値を特徴量として用いる.

$$qsim(q, c_t) = 1 - \frac{D_L(q, c_t)}{L_{max}(q, c_t)}.$$
 (2)

 $D_L(q,c_t)$ はレーベンシュタイン距離を, $L_{max}(q,c_t)$ は q と c_t のうち文字列として長い方の文字列長を表す.

4.3.3 文字数(LEN)

文字数の多いコメントには、動画の内容とは無関係なコメントが多いため、文字数が少ないほど有用度は高いと仮定し、 c_t の文字数の逆数を特徴量として用いる.

4.3.4 出現頻度 (FREQ)

 V_q 内で出現数の多いコメントは認知度の高いコメントであり有用度は高いと仮定し、 $\log_{10} freq(c)$ を特徴量として用いる. freq(c) は V_q 内での c の総出現数を表す.

4.3.5 正規化前のコメントのバリエーション (VAR)

 V_q 内の動画に投稿されたコメントの中で,正規化すると c になるコメントの種類数が多いほど,表現に多様なバリエーションを持つコメントであると言える。 バリエーションの多いコメントほど有用であると仮定し,c の正規化前のコメントの種類数を特徴量として用いる。

4.3.6 形容詞の有無(ADJ)

形容詞を含むコメントの有用度は高いと仮定し、 c_t に形容詞が含まれていれば 1 、含まれていなければ 0 とする特徴量を用いる。形容詞の有無の判別には日本語形態素解析エンジン MeCab [23] を用いた。Yamamoto ら [13] は動画の印象を推定するうえで形容詞の有無は有用な特徴であることを報告している。

4.3.7 形容動詞の有無(VADJ)

形容動詞を含むコメントの有用度は高いと仮定し、 c_t に 形容動詞が含まれていれば 1 、含まれていなければ 0 とする特徴量を用いる。Yamamoto ら [13] は動画の印象を推定するうえで形容動詞の有無は有用な特徴であることを報告している。

4.3.8 エントロピー (ENT)

ニコニコ動画では、拍手を表す「888」というコメントが動画の最後に集中的に投稿されるように、動画内の特定のタイミングに偏って投稿されるコメントが存在する。そのようなコメントは形式的なコメントであり有用度は低く、動画の再生時刻に依存せず満遍なく分布しているコメントは動画の内容に反応して投稿されたコメントであり有用度は高いと仮定する。この仮定に基づき、 V_q の各動画の長さを 20 分割したうえで、動画ごとに c が動画中で出現する

タイミングとその頻度を表すヒストグラムを作成し、その 後ひとつのヒストグラムに合成する. 合成されたヒストグ ラムのエントロピーの値を特徴量として用いる.

4.3.9 サビ区間での出現確率 (CHORUS)

Yamamoto ら [13] は、楽曲のサビ区間に投稿されたコメントは楽曲の印象の推定に有用であることを報告している。本研究でも、サビ区間に投稿される確率の高いコメントの有用度は高いと仮定し、 V_q 内でのコメント c の総出現数に対する、 V_q 内でのサビ区間の c の出現数の割合を特徴量として用いる。楽曲のサビ区間の検出には Goto [24] により提案された手法を用いる。

4.3.10 類似コメントの種類数 (SNUM)

c と類似したコメントは c と同様の内容を表しており、類似コメントが多いほど c が言及している内容は重要、つまり c_t の有用度は高いと仮定する。本研究では、c と最初の 2 文字が一致し、c との正規化レーベンシュタイン距離が 0.4 以下の正規化コメントを c の類似コメントと定義し、 V_q に投稿された c の類似コメントの種類数を特徴量として用いる。以下、c とその類似コメントの集合を S_c で表す。

4.3.11 類似コメントを考慮した出現頻度(SFREQ)

4.3.4 項で述べた出現頻度を S_c 内の各コメントについて求め, $\log_{10}\sum_{c\in S_c}freq(c)$ を特徴量として用いる.

4.3.12 類似コメントを考慮したエントロピー (SENT)

 S_c 内の各コメントが投稿された時刻を基に 4.3.8 項で述べたヒストグラムを作成し、そのエントロピーを特徴量として用いる.

4.3.13 類似コメントを考慮したサビ区間での出現確率 (SCHORUS)

 S_c 内の各コメントが投稿された時刻を基に 4.3.9 項で述べたサビ区間での出現確率を求め、特徴量として用いる.

4.3.14 文字バイグラム(BGR)

あるクエリにとって有用なコメントに含まれる文字列は他のクエリにとっても有用度が高いと仮定し,cの文字バイグラムを特徴に利用する.具体的には,データセットに含まれる全ての正規化コメントの文字バイグラムを抽出し,各バイグラムがcに含まれれば1,含まれなければ0とする特徴量を用いる *8.

5. 実験

5.1 クエリ

ニコニコ動画の検索ログデータは一般に公開されていないため、検索ログから実験に使用するクエリを抽出することはできない。そこで、タグとクエリで使用される語彙は類似しているという性質 [25] を利用して、本実験ではタグをクエリとして用いる。具体的には、データセットに含まれる全てのタグについて、そのタグが付与されている動画

*8 他の特徴は1次元の数値であるが,本特徴は多次元のベクトルである.

数をカウントし, その数の上位 50 個のタグを実験に使用する. 各クエリで検索される動画数の最大値は 5,032, 最小値は 117, 平均値は 546, 中央値は 309 であった.

5.2 正解データ

コメントの有用度の推定精度を評価するために,人手で正解データを作成する.実験に使用したクエリで有用度を推定する対象となる正規化コメント (C_q) の平均値は2789.5 個であり,すべてのコメントの有用度を人手でラベル付けするのは困難である.そこで,まず C_q に含まれるコメントを 4.3.1 項で述べたクエリとの関連度に基づいて 20 個のグループに分割する.このとき,たとえば C_q に含まれるコメント数が 1,000 個であった場合, 1 番目のグループには関連度が 1 位から 50 位のコメント, 2 番目のグループには関連度が 51 位から 100 位のコメント,というように分割する.次に,各グループから 10 個のコメントをランダムに選択することで,各クエリにつき 200 個のコメントを評価用にサンプリングする.

有用度の正解データの作成は以下の手順で行った. まず, 評価者にクエリを見せ,そのクエリに対してサンプリング された 200 個の評価用の正規化コメントを 4.1.2 項で述べ た方法で正規化前のコメントに復元したうえで,ランダム な順序で表示する.各コメントについて、たとえばクエリ が「初音ミク」,コメントが「爽やかー」であれば,「『初音 ミク』に関する動画で『爽やかー』と言われている動画を 検索したいですか?」と質問する. 評価者は,「検索した い」「それなりに検索したい」「検索したくない」のいずれ かを選択する. このようにして,50 クエリで合計1万件の 質問を作成した. 有用度のラベル付けは日常的にニコニコ 動画を視聴している8名の評価者により行われ*9,各コメ ントの質問に5名の評価者が回答した. すべての回答が終 了したら,「検索したい」を2点,「それなりに検索したい」 を1点、「検索したくない」を0点として5名の評価者の平 均値を求め、その値をコメントの有用度の正解値とした.

5.3 有用度推定

5.3.1 実験方法

作成した正解データを基に、SVM の線形回帰によりコメントの有用度の推定を行った。モデルの構築には SVM のライブラリである LIBSVM [26] を標準的なパラメータで使用し、各特徴の値は特徴内の最大値が 1、最小値が 0 になるように正規化を行った。有用度を推定する際は leave-one-out 交差検証、つまり実験に用いた 50 個のクエリのうち 49 個を訓練データとして使用し、残りの1 個をテストデータとして使用する操作をすべてのクエリに対して行った。評価指標には、スピアマンの順位相関係数、ケンドールのタウ、平均二乗平方根誤差 (RMSE) を用いた。ス

^{*9} 評価者の中に著者は含まれない.

表 1 選択された特徴に対するスピアマンの順位相関係数,ケンドールのタウ,RMSEの推移.「選択された特徴」の文字列が表す

特徴け	43	節を参照.	

特徴は 4.3 即を参照.				
ステップ	選択された特徴	スピアマン	ケンドール	RMSE
1	BGR	0.7454	0.6003	0.2779
2	LEN	0.7526	0.6075	0.2744
3	ADJ	0.7540	0.6086	0.2737
4	SFREQ	0.7546	0.6092	0.2737
5	VADJ	0.7547	0.6092	0.2736
6	QSIM	0.7547	0.6094	0.2735
7	SENT	0.7552	0.6099	0.2735
8	ENT	0.7555	0.6102	0.2735
9	CHORUS	0.7554	0.6100	0.2735
10	QREL	0.7556	0.6103	0.2735
11	SCHORUS	0.7555	0.6102	0.2735
12	VAR	0.7562	0.6109	0.2735
13	FREQ	0.7561	0.6109	0.2735
14	SNUM	0.7561	0.6109	0.2735

ピアマンの順位相関係数およびケンドールのタウは [-1,1] の値を取り、値が大きいほど推定精度が高いことを表す。 RMSE は値が 0 に近いほど推定精度が高いことを表す.

また,以下の逐次特徴選択により有用度の推定に重要な特徴を検証した.(1)4.3節で述べた特徴の集合を F_{all} ,SVM の線形回帰に用いる特徴の集合を F_{use} とし, $F_{use} = \phi$ で初期化する.(2) F_{all} 内の各特徴を F_{use} に追加した際の全クエリでの RMSE の平均値を求め,その値が最小となる特徴 f を求める.(3)f を F_{all} から F_{use} に移し, $F_{all} = \phi$ であれば終了, $F_{all} \neq \phi$ であれば(2)に戻る.

5.3.2 結果

5.3.1 項の手順で特徴の選択を行った際の各評価指標の全クエリの平均値の推移を表1に示す.この結果を見ると、コメントの有用度を推定するうえで文字バイグラムが非常に有用な特徴であることがわかる.つまり、あるクエリにおいて有用なコメントは他のクエリにおいても有用であることを示している。文字バイグラムに加えて文字数、形容詞の有無、類似コメントを考慮した出現頻度、形容動詞の有無、クエリ類似度を用いたときにRMSEの値は最小となり、それ以上特徴を加えてもRMSEの値に大きな変化は見られなかった。RMSEが最小となったときのスピアマンの順位相関係数の値は0.7547であり、高い精度で有用度の推定ができていることが示された.

次に、RMSEの値が最小となった特徴の集合を用いた際の、クエリごとのスピアマンの順位相関係数を図2に示す、大部分のクエリでは相関係数が0.6以上であったが、「踊ってみた」と「合唱シリーズ」の2つのクエリは他のクエリに比べて相関係数は低かった。この理由として、これらのクエリに関する動画には、ユーザに対するコメントであることを明示するためにユーザの名前を含むコメントが多いことがあげられる。たとえば「踊ってみた」では「まぁむ」というユーザに対する「まぁむかわええ」というコメントの有用度は「かわええ」という文字列の影響で1.408と高

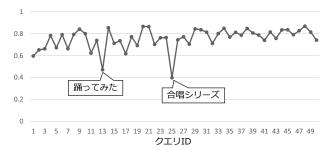


図 2 クエリごとのスピアマンの順位相関係数.

表 2 クエリ「音楽」における、コメントの有用度の推定結果と正解

順位	推	定結果	正解データ		
1	調教神すぐる	(1.6)	これはいい調教 (2.0)		
2	絵&歌詞神すぎる	(1.6)	ヘッドホン推奨(1.8)		
3	ピアノやべえ	(1.8)	聞き入ってしまう(1.8)		
4	これはいい調教	(2.0)	ピアノやべえ(1.8)		
5	やばい鳥肌	(1.2)	マイリス直行(1.8)		
6	レンリン最高	(1.0)	良い歌(1.6)		
7	あなたが神か	(1.4)	これはいい PV(1.6)		
8	まろやか野菜ジュース	(0.6)	CD ほしい (1.6)		
9	これは楽しい	(1.2)	ミリオンおめー(1.6)		
10	これはすごかった	(1.2)	調教神すぐる(1.6)		
11	処女作・・・だと・・・	(1.4)	絵&歌詞神すぎる(1.6)		

く推定されており、200 コメント中 6 位であった. しかし、評価者は特定のユーザに対して向けられたコメントは有用ではないと判断し、有用度の正解値は 0.4 と低かった. Web 上には踊り手や歌い手のユーザ名の一覧ページが存在するため、その情報を基にコメント中のユーザ名の有無を特徴量とすることが解決策のひとつとしてあげられる.

最後に、「音楽」というクエリについて、正解値が 1.6 以上の 11 件のコメントと、推定された有用度の上位 11 件のコメントの比較結果を表 2 に示す。正解データの上位 11 件のうち 4 件のコメントは推定結果の上位 11 件に含まれており、また推定結果の上位 11 件中 10 件のコメントの正解値は 1.0 以上であった。ユーザにコメントを提示する際は、ノイズとなるコメントの数を少なくすることが重要であるため、「音楽」というクエリの推定結果は本研究の目的のもとでは望ましい結果である。

5.4 検索結果の重複度

コメントに基づいて動画のランキングを生成した際に、動画の再生数やコメント数に基づく従来のランキングと大部分の動画が重複していれば、ユーザが新しい動画に出会うためのランキングを生成するという観点における有用性は低い. そこで本節では、コメントに基づく動画のランキングと、従来のランキングの重複度を検証する.

5.4.1 実験方法

本実験では、5.3 節の実験で RMSE の値が最小となった 特徴の集合 (BGR, LEN, ADJ, SFREQ, VADJ, QSIM) によって推定されたコメントの有用度を用いる。各クエリ

表 3 コメントに基づく動画のランキング結果の上位 k 件と,従来の指標に基づく動画のランキング結果の上位 k 件の重複動画数の全クエリの平均値.括弧内の値は標準偏差を表す.

ランキング指標	上位 5 件	上位 10 件	上位 20 件	上位 30 件
再生数	0.04703	0.3303	1.218	2.923
丹生奴	(0.2214)	(0.6852)	(1.525)	(3.064)
マイリスト数	0.08384	0.3348	1.316	2.958
マイックト奴	(0.2847)	(0.6422)	(1.531)	(2.897)
コメント数	0.03272	0.2494	1.032	2.491
コクンド奴	(0.1781)	(0.5642)	(1.413)	(2.677)

で推定された有用度の上位 10 コメント,計 500 コメントを評価に使用し,クエリq とコメントc によって生成される動画のランキングを以下の方法で評価した.(1) V_q の中でc およびその類似コメントの集合 S_c に含まれるコメントが 1 件以上投稿された動画集合を求め,その投稿数が多い順に動画をランキングする.これを R_1 とする.(2) V_q 内の動画を,再生数の多い順にランキングした結果を R_2 とする.(3) R_1 と R_2 の上位 k 件で重複している動画の数を求める.本実験では k=5,10,20,30 とする.(1) から(3) の操作を,再生数の他にマイリスト数,コメント数の多い順にランキングした結果に対しても行う.

5.4.2 結果

再生数,マイリスト数,コメント数を対象としたときの,重複度の全クエリでの平均値を表3に示す.いずれの指標においても,コメントに基づくランキングとの重複度は低く,上位30件まで見た場合でも重複する動画数は3件未満であった.この結果から,従来のランキング指標では上位に表示されない動画であっても,コメントに基づくランキングを行うことで上位に表示されることが明らかになった.つまり,ユーザがコメントに基づいた自分好みの動画ランキングを生成し,従来の検索では発見が困難であった動画の視聴が可能になることを示している.

6. SmartVideoRanking

5章の実験の結果を基に、コメントに基づく動画ランキングシステム SmartVideoRanking を実装した。SmartVideoRanking はまず5章の実験で使用した50個のクエリをユーザに提示する(図3上部)*10. たとえばユーザが「初音ミク」というクエリを選択すると、SmartVideoRankingはクエリに関する動画に投稿されたコメントの中でランキング指標として有用度の高い50個のコメントをユーザに提示する。50個のコメントを選択する際は、有用度の高いコメントから順に選択し、選択されたコメントの類似コメントは選択しないようにすることで多様化を行っている。各コメントはクエリをタグに持つ動画の中で、そのコメントおよび類似コメントが最も多く投稿された動画のサムネイルと共に表示され、有用度の高いコメントほど大きく表

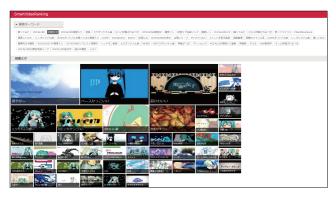


図 3 クエリ「初音ミク」選択時の結果画面.



図 4 コメント「かっこいいな」選択時の結果画面.

示される(図3下部). これによって、クエリに関する知識のないユーザであっても、まずは大きく表示された代表的なコメントを選択してランキングを生成するようにインタラクションを促すことができる.

ユーザが「かっこいいな」というコメントに興味を持ち 選択をすると、SmartVideoRanking は図 4 に示した結果 画面を表示する. 右下部には「かっこいいな」という指標 に基づいた動画のランキングが生成される. 動画のランキ ングは、選択したコメントおよびその類似コメントが多く 投稿されている動画ほど上位に表示されるように生成され る. 右上部には動画プレイヤーが表示され, デフォルトで は生成されたランキングの1位の動画が表示される. 動画 プレイヤーの下にはその動画における「かっこいいな」と その類似コメントの投稿時刻と出現頻度が折れ線グラフで 表示される.この折れ線グラフはクリックが可能であり, クリックをした位置に対応する動画の再生時刻にジャンプ することができる. これにより, たとえば折れ線グラフ上 で最もコメント数の多い箇所をクリックすることで動画の 中でも「かっこいいな」という反応の大きいシーンの視聴 などが可能になる. また, ランキング内の動画を選択する ことで、ユーザは「かっこいいな」という反応が大きい様々 な動画を視聴できる.

ユーザが選択しなかったコメントは,選択したコメント との関連度が高い順に並び替えられて表示され,関連度が 閾値未満のサムネイルは薄く表示される(図4左部). コメ

^{*10} 本稿では処理時間の問題から、事前に処理を終えたクエリのみ提示しているが、手法上はクエリに制限はなく、ユーザが自由にクエリを入力できるよう拡張可能である.

ント c_1 と c_2 の関連度は $\frac{|V_{q,c_1} \cap V_{q,c_2}|}{\sqrt{|V_{q,c_1}||V_{q,c_2}|}}$ により求め,閾値は 0.2 とした. $V_{q,c}$ は V_q の中で S_c 内のコメントが1件以上投稿されている動画集合を表す.これにより,ユーザは興味のあるコメントと関連のあるコメントを次々と選択して自分好みの動画ランキングを生成できる.関連度の低いコメントであっても興味があればそのコメントを指標としたランキングを生成できる.このように Smart Video Rankingを使用することで,ユーザは自分好みの動画ランキングを生成し,多様な動画の閲覧が可能になる *11.

7. まとめ

本稿では、視聴者の反応に基づく動画ランキングシステム Smart VideoRanking を提案した.動画に投稿された時刻同期コメントは視聴者の反応を抽出する情報源として適しているという仮定のもと、SVM の線形回帰を用いてランキングの指標としてのコメントの有用度を推定した.50個のクエリを用いた実験の結果、推定された有用度と評価者によって判定された有用度の相関は平均で0.7547を達成した.また、再生数やマイリスト数などの従来のランキングでは上位に出現しない動画であっても、視聴者の反応に基づくランキングでは上位に出現し、ユーザがより多様な動画を閲覧できることを示した.

今後の課題のひとつは、提案したシステムをユーザに使用してもらい、その有用性を定性的・定量的に評価することである。提案システムを使用した際の、ユーザが目的の動画を発見するまでの時間や、システムとのインタラクションの頻度を計測することで、コメントに基づくランキングの生成がユーザの動画検索時の振る舞いに、どのように影響を与えるかを検証する予定である。

謝辞

本研究の一部は、文部科学省科学研究費補助金研究活動スタート支援(課題番号 15H06887) および科学技術振興機構OngaCREST プロジェクトの支援を受けた.

参考文献

- M. Richardson, E. Dominowska and R. Ragno: "Predicting clicks: Estimating the click-through rate for new ads", WWW, pp. 521–530 (2007).
- [2] M. Cha *et al.*: "I tube, you tube, everybody tubes: Analyzing the world's largest user generated content video system", IMC, pp. 1–14 (2007).
- [3] B. Liu, M. Hu and J. Cheng: "Opinion observer: Analyzing and comparing opinions on the web", WWW, pp. 342–351 (2005).
- [4] K. Yatani *et al.*: "Review spotlight: A user interface for summarizing user-generated reviews using adjectivenoun word pairs", CHI, pp. 1541–1550 (2011).
- [5] 後藤真孝: "CGM の現在と未来: 初音ミク, ニコニコ動画, ピアプロの切り拓いた世界:1. 初音ミク, ニコニコ動画, ピアプロが切り拓いた CGM 現象", 情報処理学会

- 論文誌, **53**, 5, pp. 466-471 (2012).
- [6] M. Hamasaki, H. Takeda and T. Nishimura: "Network analysis of massively collaborative creation of multimedia contents: Case study of hatsune miku videos on nico nico douga", UXTV, pp. 165–168 (2008).
- [7] C. Cayari: "The youtube effect: How youtube has provided new ways to consume, create, and share music", International Journal of Education & the Arts, 12, pp. 1–28 (2011).
- [8] 濱野智史: "CGM の現在と未来: 初音ミク, ニコニコ動画, ピアプロの切り拓いた世界: 5. ニコニコ動画はいかなる点で特異なのか: 「擬似同期」「N次創作」「fluxonomy (フラクソノミー)」", 情報処理学会論文誌, **53**, 5, pp. 489–494 (2012).
- [9] Y. Wai Gen et al.: "Are web user comments useful for search?", LSDS-IR, pp. 63–70 (2009).
- [10] S. V. Chelaru, C. Orellana-Rodriguez and I. S. Altingovde: "Can social features help learning to rank youtube videos?", WISE, pp. 552–566 (2012).
- [11] S. Nakamura and K. Tanaka: "Video search by impression extracted from social annotation", WISE, pp. 401–414 (2009).
- [12] 佃洸摂他: "映像に付与されたコメントを用いた登場人物が注目されるシーンの推定", 情報処理学会論文誌, **52**, 12, pp. 3471-3482 (2011).
- [13] T. Yamamoto and S. Nakamura: "Leveraging viewer comments for mood classification of music video clips", SIGIR, pp. 797–800 (2013).
- [15] N. A. Diakopoulos and D. A. Shamma: "Characterizing debate performance via aggregated twitter sentiment", CHI, pp. 1195–1198 (2010).
- [16] V. Dang and B. W. Croft: "Query reformulation using anchor text", WSDM, pp. 41–50 (2010).
- [17] R. Jones *et al.*: "Generating query substitutions", WWW, pp. 387–396 (2006).
- [18] H. Ma et al.: "Learning latent semantic relations from clickthrough data for query suggestion", CIKM, pp. 709– 718 (2008).
- [19] Q. Mei, D. Zhou and K. Church: "Query suggestion using hitting time", CIKM, pp. 469–478 (2008).
- [20] M. Hamasaki and M. Goto: "Songrium: A music browsing assistance service based on visualization of massive open collaboration within music content creation community", WikiSym, pp. 4:1–4:10 (2013).
- [21] B. Croft, D. Metzler and T. Strohman: "Search Engines: Information Retrieval in Practice", Addison-Wesley Publishing Company, USA (2009).
- [22] L. Yujian and L. Bo: "A normalized levenshtein distance metric", Pattern Analysis and Machine Intelligence, IEEE Transactions, 29, 6, pp. 1091–1095 (2007).
- [23] T. Kudo, K. Yamamoto and Y. Matsumoto: "Applying conditional random fields to japanese morphological analysis", EMNLP, pp. 230–237 (2004).
- [24] M. Goto: "A chorus section detection method for musical audio signals and its application to a music listening station", Audio, Speech, and Language Processing, IEEE Transactions, 14, 5, pp. 1783–1794 (2006).
- [25] M. J. Carman et al.: "A statistical comparison of tag and query logs", SIGIR, pp. 123–130 (2009).
- [26] C.-C. Chang and C.-J. Lin: "LIBSVM: A library for support vector machines", ACM Transactions on Intelligent Systems and Technology, **2**, 3, pp. 27:1–27:27 (2011).

^{*11} デモ動画は http://youtu.be/pQoAeOOTsbk から視聴可能.