

ニコニコ動画におけるユーザのコメント機能を用いた動画検索支援 A method for retrieving similar videos by using comments in Niconico

天笠 忠甫

Tadasuke Amagasa

法政大学情報科学部デジタルメディア学科

E-mail: tadasuke.amagasa.rg@stu.hosei.ac.jp

Abstract

The users of video sharing services are over thirty million. Although the number of the videos increases, the availability of information about the contents of videos is still limited. Therefore, it is difficult for users to choose videos with contents that they want to see. The video sharing service called "Niconico" is popular in Japan. Unlike other services, Niconico has a unique function that allows viewers to add comments to videos during the playback. Such comments are expected to contain information about the videos. This paper proposes a method for retrieving similar videos by using comments in Niconico. The retrieval of similar videos uses the scores of feelings computed from comments. The scores of feelings consist of three axes of the pleasant/sad, delightful/angry, and peaceful/nervous feelings that range from zero to one. The method retrieves similar videos by computing the similarities based on the scores of feelings in the case of the whole videos and in the case of dividing videos into scenes. This paper shows the results of the experiments that used about one thousand six hundred videos.

1. はじめに

近年、動画共有サービスのユーザ数は 3000 万人を超え、動画の総数は増え続けている。その一方で、現在の動画共有サービスには動画内容に関する情報が少ない為、ユーザが視聴する動画を選択する際に、見たい内容がどの動画に含まれるか判断することは難しくなっている問題がある。

日本ではニコニコ動画[1]という動画共有サービスの人気が高い。ニコニコ動画は、他の動画共有サービスと違い、動画の再生中にコメントを付加できるというユニークな機能を持つ。このようなコメントには動画内容に関する情報が含まれていると考えられる。しかし、ニコニコ動画で動画を検索する際に参照されているのは、動画のタイトル、動画の情報となるタグ、再生回数、コメント数、マイリスト登録数、動画の再生数である。この為、ニコニコ動画の特徴であるコメント機能に対しての検索機能は限られている。

本研究ではこのニコニコ動画のコメントを利用した類似動画の検索手法を提案する。本研究では特にコメント

に含まれている感情を利用することに主眼を置き、コメントから感情のスコアを算出して類似動画の検索に用いる。感情のスコアは「楽しい⇔悲しい」、「うれしい⇔怒り」、「のどか⇔緊迫」の 3 軸で求められ、対照的なポジティブまたはネガティブの感情が 1~0 の範囲で算出される。本研究ではこのような感情のスコアを用いて、動画全体の場合、またはシーンに区切った場合の動画間の類似度を求め、類似動画の検索を実現する。本研究ではこの検索手法を実装した上で、約 1600 件の動画を対象に実験を行った。

2. 関連研究

関連研究として、特に以下のものを参考にした。

2.1. 新聞記事を対象とするテキスト印象マイニング手法の設計と評価

熊本ら[2]は、新聞記事を例として取り上げ、記事を読んだ人々が感じる印象を記事そのものから抽出するテキスト印象マイニング手法を提案している。新聞記事データベースを解析し、記事に現れる各単語が記事の印象に及ぼす影響を数値化した印象辞書を構築するとともに、この印象辞書を用いて記事の印象値を算出する手法を開発している。更に、この算出法が記事から算出する印象値と人々がその記事を読んだときに感じる印象値との対応関係を回帰分析により調べ、その結果得られる回帰式を用いて算出した印象値を補正するという方法で高精度なテキスト印象マイニングを実現している。提案手法の有効性を検証する為に行った被験者実験では、それぞれの印象尺度における平均誤差が 0.69, 0.49, 0.64 となり、特に「うれしい⇔怒り」に対しては高い精度を得ている。本研究ではここで開発された感情辞書と感情値算出ツール[3]を利用している。

2.2. 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

佃ら[4]は、視聴者の反応として映像の再生時刻に沿って付与されたコメントを用いて、映像に登場する人物が視聴者の注目を集めているシーンの推定と各シーンにおける各登場人物の活躍の度合いの推定を行う手法を提案している。また、登場人物の名前を含むコメントの数とその周辺の感情を含むコメントの数からその登場人物に対する注目の大きさを求めることでダイジェスト映像を抽出している。本研究ではニコニコ動画のコメントデー

タの利用，シーン毎の比較，コメントデータからの感情の抽出などにおいて仮らの研究を参考にしている。

2.3. 視聴者コメントを用いた広告動画挿入タイミング決定アルゴリズムの提案と評価

齋藤ら[5]は，動画の視聴を妨げないタイミングで広告動画を挿入する為，視聴者コメントを用いたミッドロール型広告動画挿入タイミング決定アルゴリズムを提案している。単位時間毎のコメント数に着目し，動画の視聴を妨げないタイミングを検出し，広告動画の挿入を試みている。さらに，広告動画挿入タイミング決定アルゴリズムを作成する為の予備実験を行い，単位時間毎のコメント数と適切な広告動画挿入タイミングの関係について調査した。さらに，予備実験により得られた知見から，広告動画挿入タイミング決定アルゴリズムを考案し，提案アルゴリズムの適用範囲について評価を行っている。本研究では取り入れなかったが，このアルゴリズムを使用することで，類似動画検索の精度を向上できる可能性がある。

3. 準備

3.1. ニコニコ動画のデータ

本研究ではニコニコ動画のデータとしてニコニコデータセット[6]を利用した。これはニコニコ動画に 2012 年 11 月初旬までに投稿された約 830 万件の動画のメタデータとそれに対するコメントデータからなっている。

動画のメタデータには，動画 ID，スレッド ID，動画タイトル，動画説明文，サムネイル画像の URL，投稿日時，動画再生数，動画フォーマット，高画質動画のファイルサイズ，低画質動画のファイルサイズ，動画の再生数，コメント数，マイリスト登録数，最近のコメントのサマリ，タグの情報が含まれており，動画データ本体は含まれていない。コメントデータには，コメント投稿日時，コメント番号，コメント書き込み再生位置，コメント本文，コマンドの情報が含まれている。これらのデータは全て JSON 形式になっている。

3.2. 感情辞書

コメントデータから感情のスコアを算出する為，感情辞書と感情値算出ツール[3]を使用する。感情辞書とは，単語とその単語が表す感情との対応関係を表すものである。この感情辞書に基づいて，テキスト(本研究ではニコニコ動画のコメント)の感情値が算出される。テキストの感情値は，テキストに出現した各単語の感情値の平均で算出される。感情算出ツールは Linux 用に作られている。文章中の繋ぎ目の処理には形態素解析ツール JUMAN [7]が利用されている。

感情軸の設定は 3 軸で行われ，「楽しい⇔悲しい」，「うれしい⇔怒り」，「のどか⇔緊迫」となっており，単語の感情値が 1 に近い場合，単語の感情が「楽しい」，「うれしい」，「のどか」に寄り(例：初受賞，勝つ，ひなまつり)，単語の感情値が 0 に近い場合，単語の感情が「悲しい」，「怒り」，「緊迫」に寄る(例：偽装，死刑，拘束する)ように算出される(表 1)。

表 1 感情辞書のスコア[3]

エントリ 単語	感情軸 1: 楽しい⇔ 悲しい	感情軸 2: うれしい⇔ 怒り	感情軸 3: のどか⇔ 緊迫
初受賞	0.862	1.000	0.808
勝つ	0.706	0.780	0.187
ひなまつり	0.847	1.000	0.977
偽装	0.245	0.075	0.297
死刑	0.051	0.081	0.295
拘束する	0.059	0.103	0.000

4. 提案手法

本研究では，3 節で説明したニコニコデータセットと感情値算出ツールを使用し，ニコニコ動画のコメントを用いた類似動画の検索を行う。

本研究では，ユーザが動画を視聴した際にその感想をコメントに書き込んでいると仮定する。そのコメントから感情を読み取って利用することで，動画間の感情のスコアが似ているものを類似動画として提示する。

動画より感情のスコアを算出する方法として，動画全体のコメントの感情のスコアと，30 秒で 1 シーンとして区切ったときのシーン毎の感情のスコアを求める。しかし，シーンの長さを設定してしまうと，動画によって再生長が違うので，検索するに当たって感情のスコアの次元が変わってしまう問題が生じる。そこで本手法では，動画の再生長，シーンの数によって左右されないように，類似動画を検索する動画を選択した際に，全ての動画において，元の動画と感情のスコアの次元を同じにすることで，この問題を解決する。スコアが導かれたら，情報検索としてユークリッド距離を用い類似度を算出する。感情のスコアは前述したように，動画全体とシーン毎に求めており，感情は 3 軸で算出されている為，1 つの空間につき 3 次元となっている。

シーンに区切った場合の類似度の計算方法の詳細を述べる，検索の元となるシーン数 5 の動画 A と，検索対象となるシーン数 3 の動画 B の類似度を求める場合，元の動画 A は 15 次元となる。それに対し検索対象の動画 B は 9 次元となってしまう。本手法では元の動画と次元を同じにする為，この場合動画 B の次元を増やすことで，計算を行っている。具体的には，動画 A のシーン A0~A4 に対し，B のシーン B0~B2 のうちの B0 と B1 を 2 回使用し，B2 を 1 回使用するようにしている。式(1)は，簡単のため感情が 1 軸であるとした場合の動画 A，B の類似度の計算方法を示したものである。このとき，動画 A は 5 次元，動画 B は 3 次元となる。動画 B のシーン B0，B1 を 2 回ずつ用いることで，動画 A，B の次元を合わせるようにしている。

$$\sqrt{\{(A0 - B0)^2 + (A1 - B0)^2 + (A2 - B1)^2 + (A3 - B1)^2 + (A4 - B2)^2\}} \quad (1)$$

このようにシーン毎の感情のスコアより求めた類似度，動画全体の感情のスコアより求めた類似度，そしてこれら 2 つの類似度を合わせたものの 3 種類の類似度を求め，距離に近い順に類似動画としている。

5. 実装

提案手法に基づき、Java で実装を行った。実装の手順を以下に記述する。

5.1. ニコニコ動画のデータ変換

ニコニコデータセットから取得した動画のメタデータとコメントデータは全て JSON 形式になっている為、これらを Java オブジェクトに変換するプログラムを作成した。JSON 形式を扱う為に Google GSON [8]ライブラリを使用した。全データは必要ない為、本研究では、動画のメタデータからは動画 ID、動画タイトル、サムネイル画像の URL のみを抽出し、コメントデータからはコメント本文のみを抽出した。

5.2. 感情のスコアの算出

次に感情値算出ツールを使用し、5.1 節で抽出したコメントデータから感情のスコアを算出した。感情値算出ツールは、Linux 上で実行した。感情のスコアは動画全体とシーン毎の 2 種類を算出している。本研究では 1 シーンの長さを 30 秒と設定している。これにより動画全体とシーン毎のコメントに出現した各単語の感情値の平均の算出値と、回帰式による補正値を得た。

5.3. GUI の作成

類似動画検索をより使いやすくする為に、Java の Swing と AWT を用いて GUI を作成した。この GUI には動画全体の場合、シーンに区切った場合、それら 2 つを合わせた場合の 3 種類の類似動画検索を行う為に、3 つのテキストフィールドと 3 つのボタンを用意している(図 1)。テキストフィールドに動画の ID を入力し、ボタンを押すことで、それぞれの類似動画検索ができるようになっている。

類似動画の検索結果は、動画の ID、サムネイル画像として表示されており、動画のタイトル、類似度、動画の URL に関してはサムネイル画像の隣のスクロールバーで閲覧できるようにしており、ID を入力した元の動画から類似度が最も高いものから順番に表示するようにしている(図 2)。

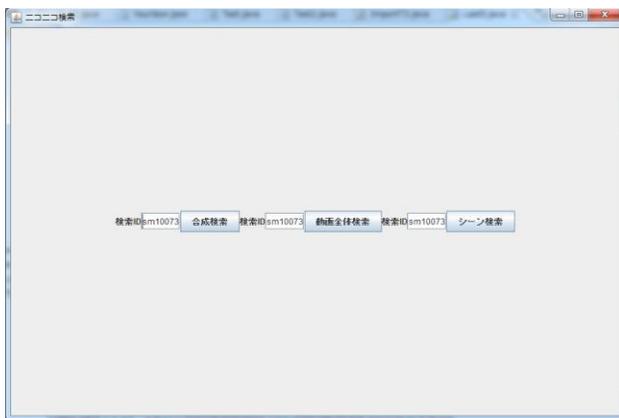


図 1 類似動画の検索画面



図 2 類似動画の検索結果

6. 実験

実験は 1616 件の動画を対象として、数件の動画を選んで類似動画を検索した上で、動画の内容、タグ情報からのジャンル、コメント内容などを比較することで行った。

6.1. 動画全体の場合

類似度が近いものから順に確認したところ、動画の内容、ジャンル、コメント内容ともに、値が近い順と元の動画を比較しても似ているとは言い難い結果であった。

6.2. シーンに区切った場合

動画の内容とコメントから比較的似ている印象を受けた。タグ情報でも同じ様なタグが見られ、動画のジャンルが同じカテゴリの物が見られることも多かった。

図 2 はゲームのオープニング動画を対象に類似動画の検索を行ったもので、最も右に並んでいる動画が検索結果である。類似度が近く上位に入る動画は、全てゲーム動画であり、コメントにも「すごい」、「かっこいい」など興奮しているようなコメントがよく見られた。

6.3. 2 つを合わせた場合

動画全体の場合に比べると元の動画に近い物が多かったが、シーン毎の場合ほどではなかった。しかし、検索対象の元の動画ではないが、類似度順に動画を並べてみると、シリーズ物の動画が並んで結果に出ている場合があることを確認した。

6.4. 実験結果のまとめ

動画全体の場合、類似動画の検索には不向きであった。一方、動画をシーンに分割することにより、動画の傾向などが再生中にタイムライン上に付加されているコメントデータから読み取ることが出来た。シーン毎に感情のスコアを算出すれば、動画全体のコメントのスコアを使用するより比較的類似している動画の検索を行うことが出来た。

7. 議論

実験の結果、シーン毎の感情のスコアを利用すること

で他の場合よりも精度が高い類似動画の検索が出来たが、満足のいく結果にはならなかった。以下では問題点を検討する。

7.1. 問題点

ニコニコ動画の最大の特徴はコメント機能ではあるが、他にも主に投稿されている動画がアニメやゲームなどのサブカルチャーが主体になっている特徴がある。アニメやゲームなどの場合、派手なシーンが多く、ユーザが入力しているコメントに興奮気味のものが多く見られた。

次にニコニコ動画には、MADムービーという、既存の音声、ゲーム、画像、動画、アニメーションなどを個人が編集、合成し、再構成したものが多い。この場合、起承転結がなく、ユーザが入力しているコメントから読み取れる感情も単一なことが多く、本研究の提案手法では扱いにくいネットスラングも多く見られた。

最後に本手法では、動画の再生長に関係なく、感情のスコアの次元を同じにさせた為、動画間の類似度が一意ではなく、元の動画に応じた値になる。感情のスコアの次元が同じもの同士であれば、関係を逆にしても類似度は変わらないが、次元が違うもの同士では、類似度(ユークリッド距離)が変わってくる。このため、検索によって得られた類似動画から、元の動画を再び類似動画として見つけることが常には可能ではないという問題がある。

7.2. 問題点の考察

7.1 節で問題点として挙げた、サブカルチャー作品が多かったり、映像作品として起承転結がなかったりすることなどに関しては、ユーザはそれを好んでニコニコ動画を利用していると考えられる為、変えることはできないだろう。本手法ではコメント情報しか参照していなかったが、ニコニコ動画に本来ある視聴数やタグなどの情報も取り入れることにより、この問題に関しては改善されると考えられる。

一方、ネットスラングに関してはこれまでも研究が行われている。石野ら[9]は、ネットスラングや特殊な固有名詞などを特徴語として捉え、その特徴語の属性を推定し、動画の要約やシーン抽出する際にその属性を推定した特徴語、属性特徴語を用いることでラベリングを行っている。このような手法を取り入れることで検索対象になる単語を増やすことができ、検索の精度は上げられると考えられる。

動画の再生長によって感情のスコアの次元が変わってしまう問題について、本研究では検索する元の動画に次元を合わせて行った。しかし、あらかじめ全ての動画の次元を同じにしておいたり、検索対象の動画の再生長を限定したりするなど、他にも対処方法が考えられる。

8. おわりに

本研究ではニコニコ動画を対象に、膨大な数の動画から類似動画を検索する手法を提案した。再生中のタイムライン上に沿って流れているコメントを利用し、3軸の感情のスコアより、動画全体の場、シーンに区切った場合、2つを合わせた場合の3種類の類似度を用いた動画検索を行った。これらの中では、シーンに区切った場

合で良い結果が得られた。しかし、ネットスラングや動画の再生長によって感情のスコアの次元が変わってしまうことなどの問題点がわかった。

今後の課題として類似動画検索の精度の向上があげられる。齋藤ら[5]は動画のシーンの区切りは映像の切り替わりの箇所であり、視聴者の動画の視聴を妨げにくいと仮定して、そこに広告を挿入するアルゴリズムを提案した。動画の視聴の妨げにならない箇所では、何らかの視聴者の感情の切り替えがあると考えられる。本研究では動画の全コメントを抽出したが、このアルゴリズムを使用してコメントの抽出位置などを決定するようにすれば、より精度の高い類似動画検索が出来ることも考えられる。

他の課題としては、より良い類似動画の検索を行う為に、視聴回数やタグ情報の参照、形態素解析のネットスラング対応を行う必要がある。そして、感情のスコアを導くにあたり、動画の再生長によって変わってくる動画間の比較をどのように行うか検討する必要がある。

文 献

- [1] ニコニコ動画:GINZA.
http://www.nicovideo.jp/video_top
- [2] 熊本忠彦,河合由紀子,田中克己,新聞記事を対象とするテキスト印象マイニング手法の設計と評価,電子情報通信学会論文誌, Vol.J94-D, No.3 (Mar.2011).
- [3] 感情辞書.
<http://www.kde.cs.tsukuba.ac.jp/~zjw/wiki/index.php?%E6%84%9F%E6%83%85%E8%BE%9E%E6%9B%B8>
- [4] 佃洗撰,中村聡史,山本岳洋,田中克己,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, Vol.52, No.12, pp.3471-3482 (Dec.2011).
- [5] 齋藤義仰,村山優子,視聴者コメントを用いた広告動画挿入タイミング決定アルゴリズムの提案と評価,情報処理学会論文誌, Vol.52, No.2, 520-528(Feb. 2011)
- [6] ニコニコデータセット.
<http://www.nii.ac.jp/cscenter/idr/nico/nico.html>
- [7] 日本語形態素解析ツール JUMAN.
<http://nlp.ist.i.kyoto-u.ac.jp/index.php?JUMAN>
- [8] Google GSON.
<http://code.google.com/p/google-gson/>
- [9] 石野克徳,折原良平,中川博之,田原康之,大須賀昭彦,フォークソノミとソーシャルアニメーションを用いた動画共有サービス利用支援の試み,情報処理学会論文誌, Vol.53, No.11, pp.2494-2506 (Nov.2012).