

印象に基づく動画検索

中村 聡史[†] 田中 克己[†]

[†] 京都大学大学院 〒606-8501 京都市左京区吉田本町

E-mail: [†] {nakamura, tanaka}@dl.kuis.kyoto-u.ac.jp

あらまし 近年、動画共有サイトが飛躍的にそのシェアを伸ばしている。ユーザは、キーワードやタグなどを利用して動画を絞り込み、人気度や新しさに基づくランキングを活用することで対象とする動画を探す。しかし、こうした動画検索は十分なものではなく、ユーザは自身の探そうとする動画にたどり着けないことが多い。例えば、笑える動画や、泣ける動画などをその度合いに応じて探すことは困難である。本稿では、ソーシャルアノテーションを時間軸に基づき印象分析し、インデックスを生成することで、印象に基づく動画検索およびランキングを可能とする。

キーワード 動画検索, リランキング, 印象

Search Video Clips based on Impression

Satoshi NAKAMURA[†] Katsumi TANAKA[†]

[†] Graduate School of Kyoto University

Yoshida-Honmachi, Sakyo-ku, Kyoto, 606-8501, Japan

E-mail: [†] {nakamura, tanaka}@dl.kuis.kyoto-u.ac.jp

Abstract Recently, video sharing Web sites have been becoming popular. User can search video clips by inputting keyword or browsing tags and rank them based on popularity or freshness. However, these methods are not enough to search target video clips. For example, it is not easy for users to find much more laughing or tear-jerker video clips. In this paper, firstly, we analyze social annotation to video clips. Then, we generate index based on each impression and realize the prototype system. Finally, we discuss the usefulness of our system.

Keyword Video search, Re-ranking, Impression

1. はじめに

近年、YouTube¹やニコニコ動画²などの動画共有サイトが飛躍的にそのシェアを伸ばしている。一方で、動画共有サイトにおける検索やランキングについては、十分な研究開発がなされていないため、多くのユーザにとって求める動画にたどり着くのは容易ではない。

動画を視聴するモチベーションの一つは気分転換である。気持ちが落ち込んでいるときに笑える動画や楽しい動画を視聴することで、落ち込んでいた気持ちを払拭したり、泣きたい、感動したいと思ったときに、泣ける動画や感動できる動画を視聴したりする。

こうした動画検索を行う場合、ユーザは動画共有サイトが提供するキーワード検索や、タグに基づく検索、人気度や新しさなどを考慮したランキングを利用する。しかし、動画をアップロードするユーザがキーワード検索を行うための十分な説明文を付けていることは少ない。また、「笑える」「泣ける」などの印象にまつわるキーワードをタイトルや説明文を付加していることはまれであるため、気分転換のための動画検索は困難である。

動画に対して「泣ける」「笑える」などの印象に関するタグが付加されていることはあるが、すべての「笑える」「泣ける」動画についてそうしたタグが付加されているわけではなく、全動画に対して微々たる量である。また、仮に印象に関するタグが

¹ <http://www.youtube.com/>

² <http://www.nicovideo.jp/>

付加されていたとしても、その印象に関する度合いはわからない。

動画のランキングについては、動画の再生数によるもの、動画に対するブックマークの多さによるもの、ユーザの投票によるもの、動画の新しさによるものなどが一般的である。動画の再生数やブックマーク数、評価などはある種のクオリティを担保するものではあるが、完全ではない。また、人気度に過ぎないという問題もある。

本研究では、動画に対するソーシャルアノテーションに注目し、感性に基づく検索を可能とする仕組みを実現する。動画に対するソーシャルアノテーションとしては、図1のようにブログやソーシャルブックマーク、コメントなどがある。本研究では特に、動画に対するコメントに注目する。

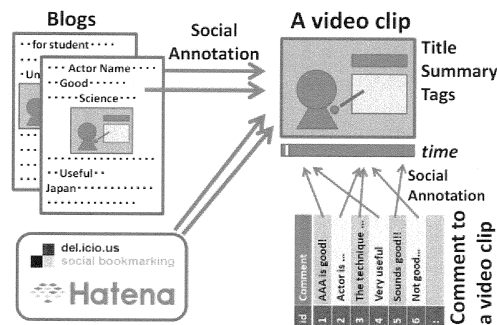


図1. 動画に対するソーシャルアノテーション

ニコニコ動画は、動画の特定の再生時間に対してコメントを付加することを可能としている動画共有サービスである。動画に対するコメントが、動画の再生時間に応じてオーバーレイ表示されるため、その場を共有する感覚で楽しむことができる。2008年12月22日の時点で191万の再生可能な動画が投稿されており、総再生数は80億、総コメント数は17億である。もっともコメントを集めている動画には1千万のコメントが投稿されている。ニコニコ動画の成功を受け、動画に対するコメント投稿サービスを提供する動画共有サイトや、マッシュアップサイトなどが増加している。我々は、今後こうしたサービスがますます増加すると考えている。

ここで我々は、動画が泣けるものである場合には悲しみに関するコメントが、笑えるものである場合に笑いに関するコメントが、また質の高い動画に対しては感謝に分類されるようなコメントが投稿されると仮定し、そのコメントの傾向に基づき動画をランキングする手法を提案する。

本稿では、まずはじめにニコニコ動画のコメントデータのソーシャルアノテーションとしての可能性を明らかにするために、コメントデータを収集し、その分析を行う。次に、印象に基づく動画インデックスとランキングシステムを提案、実装する。最後に、実際に運用することで提案手法の有用性について考察を行う。

2. 関連研究

映像検索の手法として、コンテンツの類似性を利用した検索に関する研究が行われている[1]。しかし、こうした研究はコンテンツに特化したものであり、視聴者の印象などを考慮したランキングに適用することはできない。

我々は過去の研究[2][3]において、テレビ番組のダイジェストを生成するため、テレビ番組に対する実況チャットのコメントを分析し、その喜びや悲しみの量を抽出してインデックスを作る仕組みを実現している。盛り上がるシーンを喜びと悲しみ表現の時間的な変化を利用して検出することにより、効果的なダイジェストを生成することに成功している。上原らは、テレビ番組に対する実況チャットを分析し、固有名詞を抽出することである俳優が映っているシーンなどを検出する仕組みを実現している[4]。これらの研究はチャットコメントというあまり質が高くないソーシャルアノテーションも集約することで、有用になることを示している。

また、我々は動画共有サイトの動画に対するアノテーションを分析し、動画の信頼性を図る研究に取り組んでいる[5][6]。この研究では動画に対する肯定・否定コメントを動画の再生時間軸、投稿されてからの経過時間軸という2つの軸で分析し、ユーザにグラフとして提示するものである。グラ

フを利用することにより、ある程度動画の見る価値を診断できることがわかっている。

本提案手法は上記のようなこれまでの手法を動画共有サイトに応用し、動画検索のためのインデックスとして利用するものである。

ソーシャルアノテーションを利用して検索結果を最適化する研究として山家らの研究[7]や Heymann らの研究[8]がある。これらの研究は、ソーシャルアノテーションの時間的変化や、アノテーション内の印象タグなどを活用し、ウェブ検索結果をリランキングするものであり、ソーシャルアノテーションの可能性を示唆している。

3. 分析

動画に対するコメントがソーシャルアノテーションとして活用可能なものかを確認するため、分析を実施した。分析は大規模データのクローリングによる基礎分析と、アプリケーションベースでの分析によるものの2つである。ここでは、動画に対するコメントサービスで最も成功を収めているニコニコ動画を対象とした。

3.1 基礎分析

ニコニコ動画に投稿されるコメントの傾向を分析するため、動画に対するコメントのクローリングを実施した。

クローリング期間は2008年10月下旬から2008年12月10日である。動画は10月下旬時点での動画IDの最大値が500万であることに注目し、1から500万までの数字をランダムに生成して収集対象動画を選択することとした。なお、収集内容はタイトル、説明文、タグ、コメント、再生数、コメント数、マイリスト数（お気に入り数）である。なお、コメントは各動画について最大1000件取得することとした。

収集した動画のうちサーバから削除されていない動画の収集件数は304,460件で、収集したコメント件数は56,473,136件であった。

動画に対してどの程度コメントが付加されているかを調査したものが図2である。横軸がデータセットに対する割合、縦軸が動画に対するコメントの投稿件数を示して

いる。38.5%の動画に対してコメントが100件以上投稿されており、5%の動画に対して2000件以上のコメントが投稿されていることがわかる。

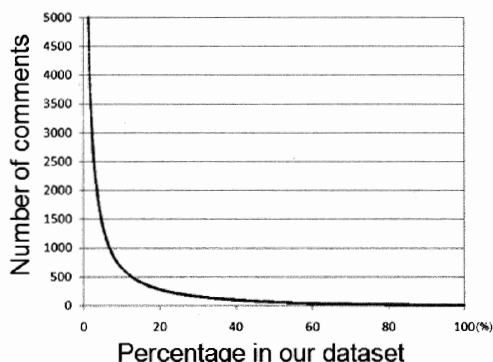


図2. 投稿コメント数とデータセット中の動画の割合との関係

ここで、動画に対するコメントの投稿数で動画をランキングした場合、どのような動画が上位にランキングされるかを調査した。調査日は2008年10月20日で、月間総合コメントランキングトップ100³、週間総合コメントランキングトップ100⁴を対象とした。実際に上位にランクされている動画を視聴すると、ただ単にキータイプするだけの動画や挨拶をするだけの動画、すでに削除されてしまった動画などであった。こうした動画に対するコメントは、実際の動画の質を担保するものではない。

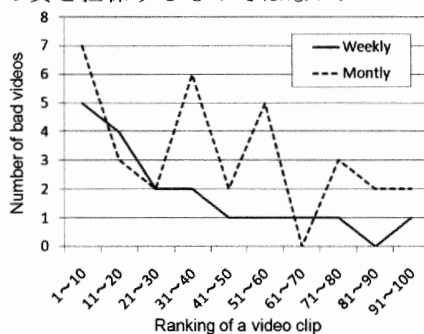


図3. コメントに基づくランキングと、無用な動画との関係性

この調査結果をグラフにしたものが図3である。横軸は動画のランキング、縦軸はコメント数が動画の質を評価するに至って

³ <http://www.nicovideo.jp/ranking/res/monthly/all>

⁴ <http://www.nicovideo.jp/ranking/res/weekly/all>

いない動画の数である。なお、ここでコメント数が動画の質を評価しているかどうかは著者らの主観に基づくものである。

結果から明らかなように、ランキングが高いほど意味のないコンテンツがランキングされている。これは、動画のランキングに単純にコメント数を使うことの危険性を示唆している。

次に、動画に対する各種印象が動画の再生時間に応じてどのように変化するかを調査した。データセットとして、ある程度コメントが付加されているものを対象とするため、データセットのうち、100件以上のコメントが付加されている動画を対象とした。対象となったデータは117,217件の動画である。また、各種印象の抽出には、我々の過去の研究[5]において実現した正規表現辞書を利用した。

ここで、各印象の抽出精度について簡易的に調査するため、10,000件のコメントをランダムに抽出し、そのコメントだけをみてどの程度網羅的にその印象に分類されるべきコメントを抽出できたか(再現率)、どの程度抽出した印象コメントの精度が良いか(適合率)を調査した。それぞれの結果は表1の通りである。Accuracyは適合率を、Coverageは再現率を表している。Happyコメント(喜びに関するコメント)の適合率はやや低い、他の印象に関しては十分な精度で抽出できていることが分かる。なお、ここではコメントのみを見て判断しているため、実際にそのコメントがどういう意図で投稿されたかは無視している。

表1. 抽出精度

	Accuracy	Coverage
Positive	95.3%	97.2%
Negative	97.1%	93.7%
Happy	85.5%	98.3%
Sad	95.8%	97.5%
Surprise	97.3%	93.5%

図4は全てのデータセットについて、動画のコメントを動画の再生時間およびそのコメントの割合でマッピングしたものである。横軸が動画の再生時間の割合で、縦軸がその動画に対して投稿された全コメントに対する各コメントの割合である。

実験結果から明らかなように、投稿されるコメントは最初に集中し、少しずつその量は減少し、最後の5%程度で再度多くのコメントが投稿されていることがわかる。また、肯定に分類されているコメントも最後に集中していることがわかる。喜びに関するコメントは、最初はあまり投稿されていないが、すぐにその数は上昇し、広く分布していることがわかる。悲しみや否定に関するコメントはほとんど投稿されていないことがわかる。

なお、このデータセットにおいて各印象のコメント数の平均をとったところ、喜びが71.24、悲しみが6.32、驚きが2.18、肯定が22.24、否定が10.25であった。

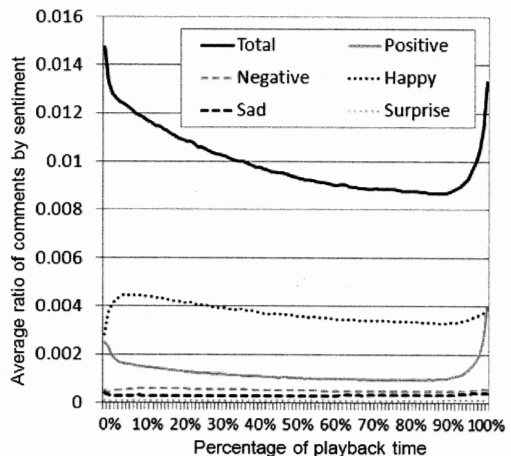


図4. 再生時間と各種感性の変化

次に、動画のタイプによって印象に関するコメントがどのような傾向で投稿されているかを調査した。ここで、分類した動画のタイプは、笑える動画集合、泣ける動画集合、質が高い動画集合、説明文と実際の動画にギャップのある釣り動画集合である。

それぞれ、笑える動画集合は「腹筋崩壊」「笑える」「吹いた」というタグが、泣ける動画集合は「涙腺崩壊」「泣ける」というタグが、質の高い動画集合は「評価されるべき」などのタグが付けられている動画、釣り動画集合は「釣り」とタグが付けられているデータセット中の動画により自動的に分類した。また、データセットの中で100件以上の投稿コメントがあるものを対象とした。

データセットから抽出された笑える動画

は 1271 件、泣ける動画は 566 件、質の高い動画は 1730 件、釣り動画は 796 件である。それぞれ、全コメント、喜びコメント、悲しみコメント、肯定コメント、否定コメントの数を動画の再生時間軸と総コメント数で正規化し平均を計算した。

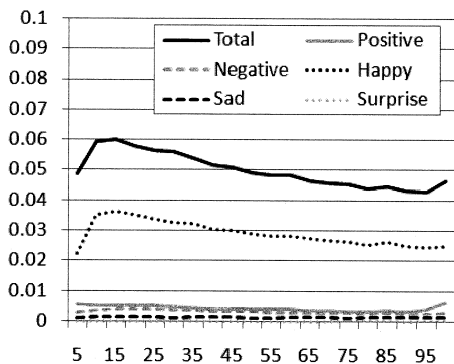


図 5. 笑える動画集合に対する各コメントの再生時間に基づく変化

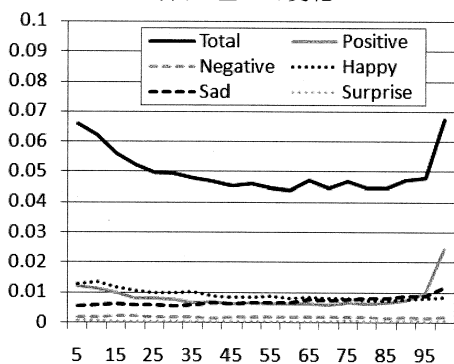


図 6. 泣ける動画集合に対する各コメントの再生時間に基づく変化

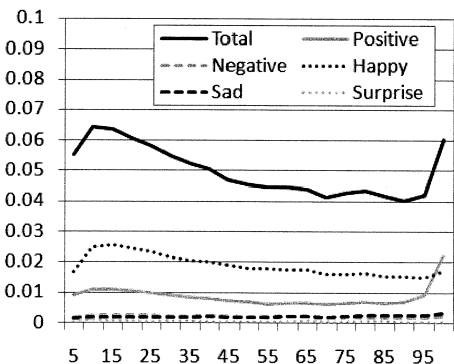


図 7. 質の高い動画集合に対する各コメントの再生時間に基づく変化

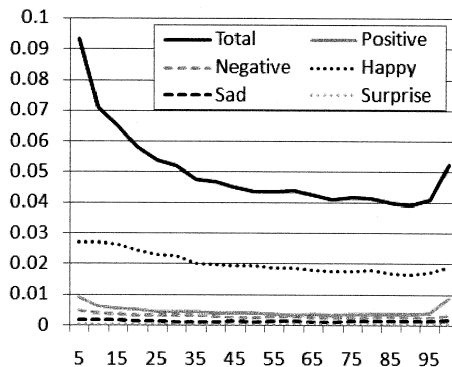


図 8. 釣り動画集合に対する各コメントの再生時間に基づく変化

実験の結果は図 5 から図 8 の通りである。図 5 は笑える動画、図 6 は泣ける動画、図 7 は質の高い動画、図 8 は釣り動画の集合に関する結果である。それぞれの図において横軸は動画の再生時間軸を 100% で表記したもの、縦軸は総コメントを 1.0 として正規化したときに含まれる各コメントの割合の平均である。

4 つの動画集合を比較するとその違いは明確である。笑える動画については、他の動画に比べ、明らかに笑いに関するコメントが多く、その量は全体コメントのうちの半数に上る。また、コメントはやや平均的に分布している。泣ける動画については、他とは異なり悲しみのコメントが強く表れている。また、動画の最後の部分に肯定コメントが集中しているのも特徴的である。質の高い動画については全体平均とあまり違いがないという結果が得られている。これは、質の高い動画には様々なタイプのものがあることが理由であり、平均化されてしまったことが原因と考えられる。釣り動画は特徴的な傾向を示しており、動画の最初にコメントが集中するという傾向がある。また、肯定コメントがほとんど現れないという傾向がある。こうした傾向を利用することで、動画の分類ができるのではと考えている。

なお、今回は動画の分類にタグだけを用いているためグラフがやや平均化されてしまっているが、タグの組み合わせや、その動画に対するお気に入りの数などを組み合わせることでより深く分類および評価することで、より明確な違いを出すことができると期待

できるため、今後詳細な分析を実施する予定である。

3.2 アプリケーションベースでの分析

動画毎に生成されるグラフは図 9～14 の通りである。このグラフにおいて、横軸は動画の再生時間、縦軸はコメント数を表している。

ここでピックアップした動画は、データセットの中で笑える動画として分類されている動画の中でコメント数が多いもの（図 9, 図 10）、泣ける動画として分類されている動画の中でコメント数が多いもの（図 11, 図 12）である。また、ニコニコ動画の中で発見した特徴的な釣り動画をピックアップしたものが（図 13, 図 14）である。

笑える動画は全般的に笑い・喜びに関するコメントが広く分布していることが分かる。一方、泣ける動画についても、悲しみに分類されるコメントが広く分布しており、特に図 12 では、最後の部分に肯定コメントが集中していることが分かる。著者らが様々な動画を視聴し観察してみたところ、質の高い泣ける動画コンテンツであればあるほど、全般的に悲しみに関するコメントが分布し、最後には動画アップロード者などに対する感謝のコメント（肯定コメント）が集中する傾向があった。

釣り動画については、多くの動画において最初に特に感情のないコメントが集中し、以後コメント数がただ減っていくという傾向を示しており、容易に検出可能である。一方、少数ではあるが図 14 のように特殊なパターンを示すこともあった。これについては、どのように発見するべきか難しい問題である。

また、図 3 に示すようなコメントの量が動画の質を示していないような動画についてアプリケーションベースで分析したところ、ほとんどの動画において印象に関するコメントが全くないか、印象に関するコメントばかりであるという傾向が見られた。こうした印象に関する情報をうまく利用することで、コメントがたくさん投稿されているが質の低い動画というものを自動的に診断できるのではと考えられる。

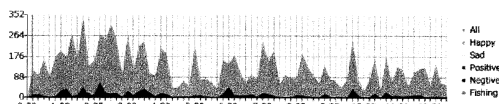


図 9. 笑える動画⁵

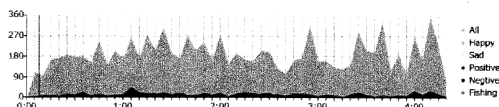


図 10. 笑える動画⁶

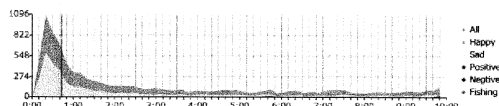


図 11. 泣ける動画⁷

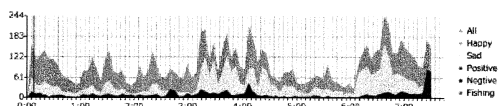


図 12. 泣ける動画⁸

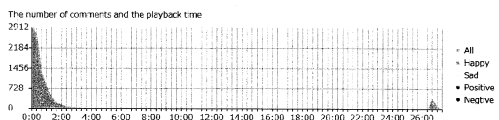


図 13. 釣り動画

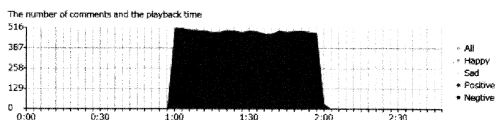


図 14. 釣り動画

4. 提案手法

印象インデックスについては、各動画について、各動画を再生時間に基づき 100 のユニットに分割し、各ユニット毎の総コメント数、喜び、悲しみ、驚き、肯定、否定コメント数を計算してデータベースに格納することで実現する。なお、各印象の抽出については、[5]の研究で利用した辞書を拡張して利用した。なお、喜びや悲しみに関するコメントがない動画は対象外とした。

印象に基づくランキングについては、今回は簡易的に笑える動画、泣ける動画の 2 種類にのみ対応するものを実現した。こ

⁵ <http://www.nicovideo.jp/watch/sm8005>

⁶ <http://www.nicovideo.jp/watch/sm1387453>

⁷ <http://www.nicovideo.jp/watch/nm2652570>

⁸ <http://www.nicovideo.jp/watch/sm1823364>

では2つのスコアリングを実現している。1つは単純にその印象に分類されるコメントの数によるもの。もう1つは、図4~8の結果により、最後の5%に肯定コメントが付加されている動画は質の高いものであるという特徴を考慮し、下記の式で動画のスコアを計算し、各動画についてそのスコアリングするものである。なお、下記の式は笑える度合のスコアリングに関するものであり、 $Happy_{all}$ はその動画中の喜びに関するコメント数、 $Total_{all}$ はその動画に対して投稿されたコメント数、 $Positive_i$ は、 i というユニットにおいて投稿された肯定コメント数を表している。

$$Score_H(v) = \log_{10} Happy_{all} \times \frac{Happy_{all}}{Total_{all}} + w \times \frac{\sum_{i=96}^{100} Positive_i}{Total_{all}}$$

提案システムを実現するため、辞書管理モジュール、コンテンツ管理モジュール、印象抽出モジュール、印象可視化モジュール、リランキングモジュールからなるシステムを実装した。また、各動画に対する印象情報を収集するため、クライアントサイドで動画を分析し、その情報をサーバに集積する仕組みを別途実現した。

ここでは2つのシステムを実装した。1つはランキングおよびスコアリングを検証するためのシステムで、これはC++により実装した。一方、実際にデータを収集し、単純に各印象の数に基づき再ランキングを行えるプロトタイプシステムをMozilla Firefoxの拡張としてXULおよびJavascriptを利用して実装した。なお、サーバはPHP+MySQL+Apacheにより実装し、APIで時間軸に基づく印象の遷移を取得できる仕組みを実現した。図15はプロトタイプシステムの動作例である。

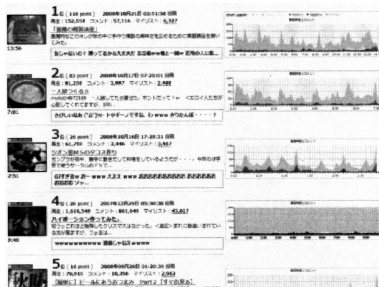


図15. システム動作例

5. 考察

複数人のユーザに依頼し、Firefoxによるプロトタイプシステムをインストールおよび利用してもらい、フィードバックを収集した。また、著者はシステムを長期にわたり利用し、その有用性を確認していった。

フィードバックによると、印象ベースでのランキングはこれまでに無いものであるということもあってか、好意的にとらえられていた。精度には多少の不満があるものの、ランキングにより100件の動画のうち必要なものを上位にランクできるという点で有用であるとの評価を得た。また、泣ける動画については結構効果的にランキングできていたが、笑える動画についてはニコニコ動画は全般的に笑いコメントが多いということもあり、それほど効果的に差をつけることができていないことがわかった。これは、笑いのポイントは人それぞれ違うため、必ずしも他人が面白いと思ったものを、本人が面白いと思うわけではないと点も大きく左右していると考えられる。

実現したプロトタイプシステムは、ランキング100件全てのグラフデータをロードするには5~10秒ほど待つ必要がある。これはサーバの性能によるものが大きいため、サーバの性能を改善するか、クライアントにキャッシュを持つことで改善できると考えられる。なお、ランキングに必要な数値データをロードするだけであれば1秒以内で動作可能であるため、数値データだけ受信してあとは随時ロードするということで対応することも考えられる。

定量的な評価実験を行っていないためランキングの性能についてはまだ評価できていない。ただ、C++で実現したシステムを利用し、いくつかのタグにより動画をリスト化したうえで、その動画を笑える度合、泣ける度合でランキングした際、単純手法より肯定コメントを考慮した手法の方が質の高い動画がランキングされている確率が高かった。ただ、これはコメント数が多い方が優先される傾向が強い単純手法の問題によるものであると考えられる。定量的な評価実験については今後実施予定である。

本稿では、動画に対するコメントについて、ユーザを分類するようなことはしていなかった。そのため、個人間での笑いや悲しみに関する感性の違いを埋めることはできていない。そこで、類似アノテーションを行うユーザグループを発見し、その情報をもとに動画をランキングする仕組みを実現する予定である。

検出精度について、評価実験ではある程度十分な精度を示しているものの、これは投稿コメントだけをチェックして精度を計測したものである。投稿されたコメントの対応する動画コンテンツを確認したわけではないため、喜びに分類されたコメントも、実際には悲しみながらのコメントであった可能性もある。ただ、実際に日々利用してみたところ問題は運用してみたところ、違和感をあまり感じないレベルであった。

本稿で利用した印象分析では、単純に喜び、悲しみ、驚き、肯定、否定、釣りというもののみを抽出した。実際には、肯定には感謝的表現が混じっていることや、その表現の対象が何なのかが明らかではないため、正確なアノテーションであるとはいえない。例えば、悲しみや喜びの表現が動画で扱われている対象に関するものなのか、発信者に対するものなのかなど不明である。今後は、そうした点についても考慮して動画に対するコメントを分析する予定である。

6. まとめと今後の課題

本稿では、まず動画に対して投稿されるコメントを分析することで、コメントが動画のインデックスとしてどこまで有用なものかを明らかにした。また、印象に基づく動画検索およびランキング手法を提案し、プロトタイプシステムを実装した。さらに、プロトタイプシステムを長期運用し、ユーザのフィードバックなどから、コメントに基づく笑える動画ランキングや、泣ける動画ランキングなどの可能性を明らかにした。今後、詳細な実験を行う必要はあるが、ある程度その笑える度合いや泣ける度合いを考慮してランキングできることも分かったことは有意義であった。

今後は、評価実験を実施することによりランキングが本当に有効なのかを定量的に診断する。また、分析結果に基づくより精度の高い印象に基づくランキング手法を実現する。一方、同一印象に分類される動画であっても、そのタイプによって投稿される印象コメントの傾向が異なると考えられるため、そうした点を考慮する予定である。また、より質の高い泣ける動画や、笑える動画などを分類およびランキングできる仕組みを実現予定である。

本手法を応用することで、類似動画の発見や、より個人の嗜好に即した動画検索およびランキングの可能性を実現できる可能性がある。そうした点についても今後取り組む予定である。

謝辞

本研究の一部は、文部科学省科学研究費補助金特定領域研究「情報爆発時代に向けた新しいIT基盤技術の研究」における、計画研究「情報爆発時代に対応するコンテンツ融合と操作環境融合に関する研究」(研究代表者：田中克己、A01-00-02、課題番号18049041)、京都大学グローバルCOEプログラム「知識循環社会のための情報学教育研究拠点」によるものです。ここに記して謝意を表すものとします。また、Mozilla Firefoxの拡張に関して実装をサポートしてくれた赤塚大典氏に深く感謝する。

文 献

- [1] N. Dimitrova, M. Abdel Mottaleb: Content-based Video Retrieval by Example Video Clip, Proc. of SPIE, Vol.3022, pp.59-70 (1997).
- [2] 宮森恒, 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 データベース Vol. 46, No. SIG 18 (TOD 28), pp.59-71 (2005年12月)
- [3] H. Miyamori, S. Nakamura, K. Tanaka: Generation of Views of TV Content Using TV Viewers' Perspectives Expressed in Live Chats on the Web, Proceedings of ACM Multimedia2005, pp.853-861 (Nov. 2005).
- [4] H. Uehara, K. Yoshida, "Annotating TV Drama based on Viewer Dialogue - Analysis of Viewers' Attention Generated on an Internet Bulletin Board -," 2005 Symposium on Applications and the Internet (SAINT'05), pp.334-340, 2005
- [5] S. Nakamura, M. Shimizu, K. Tanaka: Can Social Annotation Support Users in Evaluating the Trustworthiness of Video Clips?, Proc. of WICOW 2008, pp.59-62 (2008).
- [6] 清水誠, 中村聡史, 田中克己: アノテーションのセンチメント分析に基づく映像の信頼性判断手法, WebDB Forum 2008 (2008).
- [7] Yanbe, Y., Jatowt. A., Nakamura, S. and Tanaka, K., Can Social Bookmarking Enhance Search in the Web?, ACM IEEE Joint Conference on Digital Libraries, pp.107-116, (2007).
- [8] P. Heymann, G. Koutrika and Hector Garcia-Molina: Can social bookmarking improve web search?, Proceedings of the international conference on Web search and Web Data Mining, pp. 195-206.