

実時間と動画時間から抽出した面白い動画コメント100連発

土方 嘉徳 早川 卓弥

大阪大学基礎工学研究科

hijikata@sys.es.osaka-u.ac.jp

概要 本稿では、筆者らが開発した実時間と動画時間を用いたコメント抽出手法をニコニコ動画に適用して抽出された動画コメントを紹介する。その面白さの分類を行い、いくつかの事例を用いて面白さの解説を行う。

キーワード 面白い動画コメント, 情報抽出, ニコニコ動画, 動画時間, 実時間

1 はじめに

近年, YouTube やニコニコ動画などの動画共有サービスが広く利用されるようになってきている。中でもニコニコ動画やFC2動画では, 視聴者は動画の再生中にコメントを投稿することができ, 人気を集めている(本稿ではコメント同期型動画サイトと呼ぶ)。投稿されたコメントは投稿された動画内の時間位置が記録され, そのタイミングで表示される。他の視聴者は, これらのコメントをあたかも動画再生中に投稿されたかのように読むことができる。

投稿されたコメントの中には, 面白いと感じられるものも存在する。しかし, これまで面白いコメントを抽出して提示することは行われてこなかった。面白いコメントが存在する動画部分をユーザに提示すれば, ユーザは動画コメントを注意して閲覧することにより, それらを見逃さずに視聴できる。我々は, 動画に同期したコメントから面白い動画コメントを抽出する手法を開発してきた [1]。

上記の提案手法では, コメントのテキスト情報は用いずに, コメントに対して他のユーザが反応したコメントが存在するかどうかを調べている。これは, あるコメントが投稿された後に, 他のユーザが多く反応している場合は, その投稿は面白いコメントである可能性が高いと考えたためである。既存のコメント同期型動画サイトでは, コメントを特定してそれに対してレスポンスを入力することができない。そのため, コメントのレスポンス関係を明示的に知ることができない。

そこで, コメントのレスポンス関係の推定のため, コメントに付与された2種類の時間情報を用いた。1つはコメントの投稿日時を表す実時間情報で, もう1つはコメントの動画中における時間位置を表す動画時間情報である。一般に内容に関連するコメントは, 動画時間において近い位置に現れることが多いと考えられる。また, あるコメントに対して反応するコメントは, そのコメン

トが投稿される前には存在しないため, コメントの実時間の前後関係から, 反応するコメントの有無を判断できる。さらに, 反応するコメントの変化量を見ることで, 面白いコメントが抽出できると考えた。

本稿では, 提案した手法をニコニコ動画に適用し, 面白い動画コメントの抽出を試みた。抽出した動画コメント周辺にある動画コメントを評価者が評価し, 面白いコメントが含まれているかどうかを確認した。また, その面白さの分類と, それがなぜ面白いのかについて分析した。

本稿の構成は以下のようである。2章で提案手法を思いつくに至った背景について説明する。3章で提案手法について説明する。4章で, 提案手法をニコニコ動画に適用し発見された面白いコメントを紹介し, その面白さの分類を試みる。最後に5章でまとめを述べる。

2 技術的背景

レスポンスコメントの推定にコメントに付与された実時間と動画時間という2種類の時間情報を用いた理由を説明する。多くの動画コメントがあったとしても, それらがある他のコメントに対する反応なのか, 動画本体の内容に対する反応なのかはわからない。一般に, 動画時間で見た時に, 動画中の面白い部分やハイライトとなる部分には多くのコメントが集中する [2, 3]。我々は, コメントの集中の時間分布の特徴は, 動画内容に対するものと面白いコメントに対するものとは異なると考えた。動画内容による面白さは時間(実時間)が経過しても不変のものである。従って, 動画内容に対してコメントの集中が発生している部分は, 時間が経過しても変化しないと考えられる。反対に, 面白いコメントに対するコメントの集中は, そのきっかけとなる面白いコメントが投稿されてから現れるものである。そのため, 面白いコメントが存在する部分には実時間の経過に従って, コメント量の変化が現れることになる。

そこで我々は, 実時間情報と動画時間情報の2種類を用い, 実時間で見た時に, 各コメントの前後でコメント

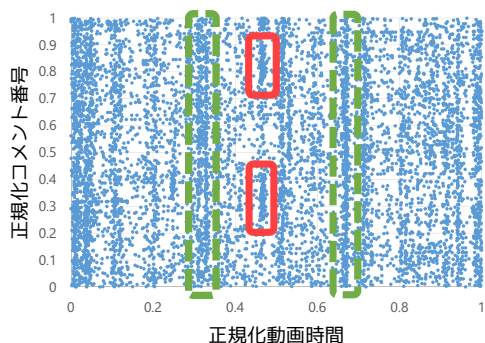


図1 コメント番号と動画時間を用いたコメントの散布図

の量が変化するかどうかを見ることにより、面白いコメントを抽出することとした。ただし、実際には実時間をそのままのスケールで用いることはせず、投稿順序のみ保管したコメント番号を利用することにした。実時間で見ると、その動画がニコニコ動画のランキングに載ったり、大手サイトで紹介されたりと、他の要因の影響を受けるからである。図1は、ある動画におけるコメントの散布図である。緑色の点線部分は動画そのものの盛り上がり、赤色の実線部分はコメントによる盛り上がりである。

3 提案手法

提案手法は、コメント番号と動画時間の二次元平面において、コメント番号の軸に沿って急激にコメント数が増えた部分を探す。そのために上記二次元平面において、あるコメントが投稿される前の周辺領域と後の周辺領域をウィンドウにより切り取る。

Algorithm 1に、提案手法のアルゴリズムを示す。アルゴリズムは、動画の全コメント及び窓の大きさ、後述する変化量計算の繰り返し回数と、全コメント数に対して面白いコメントとして抽出するコメント数の割合を入力として、面白いコメントと考えられるコメントを出力する。*CommentList*はある動画の全てのコメントを格納した配列である。この配列の要素はコメントの情報を格納した構造体であり、コメント番号、動画時間、コメント本文およびコメントの色や大きさ、位置の指定が記憶されており、さらに計算したコメント数の変化量を記憶するコメント密度差（実数型）を持つ。**Algorithm 1**の引数と内部変数であるが、 L はウィンドウの動画時間軸に対する幅である。 C はコメント A を起点として、その投稿以前と投稿以降のコメント番号の範囲を示すためのコメント数である。 $CList$ は C の大きさを格納した配列であり、コメントの変化量を N 回繰り返し計算する際に繰り返し回数に応じて C の大きさを変化させるために用意する。これについては後述する。 K は最終

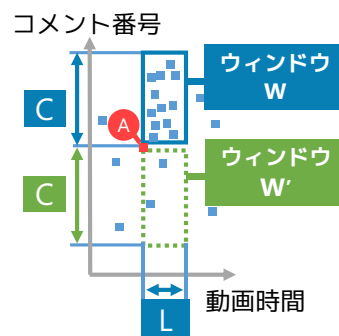


図2 ウィンドウの模式図

的に全コメント数に対して面白いコメントとして抽出するコメント数の割合を決定する変数（実数型）である。

ここでウィンドウの定義について、図2を用いて説明する。図2はウィンドウの模式図である。基準となるコメントを A とする。 A を基準として2つのウィンドウ W と W' を作る。 W と W' はそれぞれ、 A の投稿前と投稿後の周辺領域を切り取るためのウィンドウである。 W と W' の動画時間軸に対する幅はどちらも L 秒である。動画中にコメントが表示される時間3秒に加え、それに反応するコメントを投稿するまでの遅延を2秒として、今回の実装では $L=5$ 秒とした。 W と W' のコメント番号軸に対する幅はどちらも C である。

C と L で定められるウィンドウを元に、関数 *CalculateCommentDensity* ではコメントごとにウィンドウ前後のコメント数の差を計算する。この関数の処理の流れを **Algorithm 2** に示す。関数内では、引数にとったコメント群に含まれるコメントの一つ一つについて次のような処理を行う。まず、対象とするコメント A を基準として、ウィンドウ W と W' 内に含まれるコメントの数をそれぞれ計算する。*CountComment* は、 C と L 、そして対象とするコメントの配列とウィンドウの起点となるコメントを入力として受け取り、ウィンドウ内のコメント数を数える関数である。引数を $-C$ とした場合、対象となるコメントのコメント番号より若いコメントを上から順に C 件見ることとなる。つまり、図2におけるウィンドウ W' 内のコメントを数えることに相当する。次に、得られたコメントの数を C で除して正規化する。本稿では、ここで得られた値をコメント密度と呼ぶ。コメントごとに、 W と W' 内のコメント密度の差を計算する。

CalculateCommentDensity の処理を、配列 $CList$ に記録された C の値ごとに繰り返し、それぞれのコメント密度の差は、コメント情報の密度差の変数に加算し記録しておく。直近のコメント量の変化のみに注目すると、偶然起きたコメントの増減によってノイズが多くな

Algorithm 1 提案手法のアルゴリズム (Main)

```

1:  $arguments(CommentList, CList, L, K, N)$ 
2: for  $i = 0$  to  $N$  do
3:    $C = CList[i]$ 
4:    $CalculateCommentDensity$ 
      ( $CommentList, C, L$ )
5: end for
6:  $ValuableCommentList =$ 
       $Extract(CommentList, K)$ 

```

Algorithm 2 提案手法のアルゴリズム (CalculateCommentDensity)

```

1:  $arguments(CommentList, C, L)$ 
2:  $Length = CommentList$  の要素数
3: for  $i = 0$  to  $Length$  do
4:    $A = CommentList[i]$ 
5:    $countW = CountComment$ 
      ( $CommentList, C, L, A$ )
6:    $countW' = CountComment$ 
      ( $CommentList, -C, L, A$ )
7:    $densityW = countW/C$ 
8:    $densityW' = countW'/C$ 
9:    $CommentList[i].density$ 
       $+ = densityW - densityW'$ 
10: end for

```

る。そこで、窓の大きさを大きくしながら毎回の密度差を加算していくことにより、コメント量の変化が持続している部分の密度差がより大きくなるようにしている。本実装では $CList = (100, 250, 500, 750, 1000)$ とした。

最後に $Extract$ 関数で、コメント密度の差が大きいコメントから上位 $K\%$ を取り出す。 $Extract$ は具体的には、 $CommentList$ をコメント密度差の大きさに降順にソートし、上位 $K\%$ のコメントを返す関数である。

4 提案手法で抽出したコメントの質的評価

4.1 評価方法

本研究では、我々の先行研究で開発した（前節で説明した）面白いコメントの抽出手法により、どのような種類の面白いコメントが抽出できるのかを質的に評価する。

2014年8月1日から3日の間にニコニコ動画総合ランキング（24時間）に掲載された動画の上位100件のうち（動画のコメント数が2000件未満のものは対象外とした）、ランダムに8個を選んだ。これらの動画に付与された動画コメントに対して提案手法を適用しコメントを抽出した。コメントを面白いと思うかどうかは人によって異なると考えたため、この評価には5人の評価者

を用いた。この5人の内訳は、ニコニコ動画を普段からよく視聴する者が3名と、ニコニコ動画をほとんど視聴しない者2名である。

評価者にはそれぞれの動画について、まずコメント無しでその動画を通して視聴してもらった。その後、提案手法によって抽出したコメントを評価してもらった。ただし、評価者に評価してもらうコメントは、提案手法が直接に抽出したコメントに加えて、そのコメントから一定範囲内（今回は動画時間で前後5秒以内、コメント番号で前後500件以内）にあるコメントも評価してもらった。評価者には、両者のコメントの区別は提示していない。

4.2 面白いコメントのタイプ分け

あるコメントに対して、評価者5人のうち少なくとも1人が面白いと感じ、その面白さを他の4人に説明したときに、それについて同調する者が少なくとも1人存在した場合に、そのコメントを面白いコメントとみなした。抽出した面白いコメント群のうち、面白いコメントと認められたものについて、上記評価者5名により、その面白さのタイプ分けをした。

何回かのタイプ分けを繰り返したところ、面白いコメントは、「ボケ」、「ツッコミ」、「定型句」、「空耳」、「職人」、「他作品への言及」の6種類があることが分かった。「ボケ」は、動画に対して新たな見方を与えるようなジョークである。「ツッコミ」は、動画の内容（たいていは理不尽な内容やジョーク）に対して意見を述べるものである。「定型句」は、ニコニコ動画の文化で特有の定型的なコメントである。「空耳」は、動画中で、本当はそのように話していない（歌っていない）のに、聞きようによっては違って聞こえた内容である。「職人」は、動画を背景にしてそこに装飾的に付与されたコメントである（一般的にこのような装飾的なコメントを行う人は「コメント職人」と呼ばれている）。「他作品への言及」とは、当該動画とは異なる作品との類似点を見つけ、その作品について言及したものである。

4.3 面白いコメントの事例紹介

抽出された面白いコメントの例と、その面白さの解説を表1に示す。

面白いコメントのタイプのうち「ボケ」は比較的、その面白さを理解しやすいものである。表1の例では、本来なら対象動画（中のキャラクター）とはつながりのないものを持ってくることにより意外性を出している。また、あたかも本物の声優が投稿したようなコメントを出すことにより面白さを出している。「ツッコミ」も、その面白さを理解しやすいものである。表1の例では、単純に感想を述べるよりも、斬新な表現や大げさな表現で視聴者の笑いを誘おうとしている。

表1 発見された面白いコメントの例

タイプ	具体例	詳細説明
ボケ	“ドラえもん のび太の (桜 Trick)” “ふなっしー” “日笠「得意分野です！」” “日笠は「得意分野だ！」と言ってるぞw”	作品のタイトルにするなら語呂が良いことと、桜 Trick という同性愛をテーマにした作品に対して、子供向け作品の代表である漫画を持ってきている点。 作品中に登場した幼児の声か、あまりに子供っぽくなく、じゃみ声で人気のふなっしーの声に似ていた点。 声優の日笠陽子さんが担当するキャラクターに特徴があるため。 同上。
ツッコミ	“VTOL かよ!” “大 参 事”	馬が飛んだことに対し、垂直離着陸機の略称を持ってきている点。 アニメ (二次元) のプロモーションビデオなのに実写のシーンが挿入されていた点。ニコニコ動画の視聴者の多くは二次元の動画を好む傾向がある。
定型句	“視聴確定” “これが現実”	何の脈絡もなく男風呂のシーンが出てきたことに対してコメントを提示している点。 意欲にあふれていた漫画家の生活が急に乱れたことに対して、コメントを提示している点。
空耳	“フればばい?? オシエテクレヨーー ”	
職人	“★ ★”	登場人物の体の一部に重畳させる形式で表示している点。
他作品への言及	“セルニア伊織フレ임ハート?” “another なら死んでる”	動画中で長い名前を持つ登場人物が自己紹介したことに対して、別アニメで似たような語呂を持つキャラクター名を持ってきている点。 ちょっとした怪我しかしらないような事故に対して、別アニメ「Another」だったら簡単に死んでいることを皮肉っている点。

これらの中で、我々は「定型句」が面白いコメントになるとは考えていなかった。「視聴確定」や「これが現実」などは、ニコニコ動画でのコメントには非常によく見られるものであるが、動画中で何の脈絡もなく男風呂のシーンが出てきたところで提示されていたり、意欲にあふれていた漫画家の生活が急に乱れるシーンで出てきており、その提示のタイミングが絶妙であった。ニコニコ動画でよくみられる定型句であっても、提示する場面やタイミングによっては、面白いコメントになることが分かった。

また、「他作品への言及」はもともと想定していなかったタイプであった。「セルニア伊織フレ임ハート?」は、動画中で不自然に長い名前の登場人物が自己紹介したときに、同じような語呂を持つ他作品の登場人物を提示している点が面白いと感じられた。ただし、この面白さは該当の他作品を知っている人にとってのみ理解できるものであった。

これらの例からも、面白いコメントの特徴を発見し、簡単なルールで抽出することは困難であることが分かる。

5 おわりに

本論文では、我々が開発を行ってきた時間情報から面白いコメントを抽出する手法をニコニコ動画に適用した。そして、上記手法で抽出したコメントから発見された面白いコメントについて分析を行った。上記手法は、実時間 (コメント番号) の方向に急激にコメント数が増えている箇所を特定し、その増加の引き金となったコメントを抽出するものである。コメントの内容は見ないで時間情報だけでコメントを抽出した場合に、どれほど面白いコメントが含まれているのかを明らかにした。

抽出されたコメントを5人の評価者で評価したところ、これらコメントは6種類のタイプに分けることがで

きることが分かった。従来から面白いとされてきた「ボケ」や「ツッコミ」だけでなく、「定型コメント」や「他作品への言及」など、ニコニコ動画特有の文化が面白さを伴っていることが分かった。

今後は、面白いコメントが付きやすい動画の箇所の分析を行い、元コンテンツと面白いコメントとの関係を明らかにしたいと考えている。

参考文献

- [1] 早川卓弥, 土方嘉徳, 西田正吾: 実時間と動画時間から面白い動画コメントを抽出する手法とその適用事例, ARG 第5回 Web インテリジェンスとインタラクション研究会 (ARG W12 研究会), pp. 33-38, (2014)
- [2] 青木秀憲, 宮下芳明: ニコニコ動画における映像要約とサビ検出の試み, 情報処理学会研究報告. HCI, ヒューマンコンピュータインタラクション研究会報告, pp.37-42(2008)
- [3] 佃洗撰, 中村聡史, 田中克己: 視聴者のコメントに基づく動画検索および推薦システムの提案, 第19回インタラクティブシステムとソフトウェアに関するワークショップ (WISS2011) 論文集, pp.78-83(2011)