

## 視聴者コメントに基づく動画検索手法

村上, 直至

九州大学情報基盤研究開発センター | 九州大学物理学会情報理学コース

伊東, 栄典

九州大学物理学会情報理学コース | 九州大学情報基盤研究開発センター

<https://hdl.handle.net/2324/16815>

---

出版情報 : 電子情報通信学会2010年総合大会講演論文集. D, pp.24-24, 2010-03-19. The Institute of Electronics, Information and Communication Engineering

バージョン :

権利関係 :

# 視聴者コメントに基づく動画の検索手法

## A multimedia content retrieve method based on comments

村上 直至  
Naomichi Murakami

伊東 栄典  
Eisuke Ito

九州大学理学部物理学科情報理学コース  
Dept. of Informatics, Kyushu University.

九州大学情報基盤研究開発センター  
Research Institute for IT, Kyushu University.

### 1. はじめに

国内で人気の高い利用者投稿型の動画コンテンツサービスに YouTube とニコニコ動画[1]がある。作成者は自身で動画を投稿し、作成者によるタイトル名、投稿日、動画長が付与される。その後、視聴者は各動画へタグ付けなどを行う。ニコニコ動画では、タグ以外にも視聴回数や、マイリスト登録数等の付属情報を取得できる。ニコニコ動画では、その大きな特徴である視聴者からのコメント付与がある。これらのコメントやタグ情報は folksonomy であるため、動画の特徴付けるマイニング資源として利用可能である。本研究では、ニコニコ動画を対象として、利用者投稿型の動画コンテンツに付随するタグ・コメント(folksonomy)を用いたコンテンツ検索手法について検討した。

### 2. 音楽コンテンツの解析

2009年12月20日現在、ニコニコ動画には340万個以上の動画が投稿されている。これらの動画データすべてを分析するのは困難であるため、本研究では「音楽」カテゴリの（「音楽」タグが付与された）動画だけを対象とした。

407,279個の「音楽」カテゴリに入る動画について、メタデータを取得した。動画のメタデータとして、投稿者、投稿日、動画長、再生回数、コメント数、マイリスト登録回数を取得した。

次に、動画再生回数、コメント数、マイリスト登録数、動画長について、順位-頻度のプロットを行った。その結果、動画長以外は Lavalette の法則[2]に近い形状を示した。動画長は五分程度のもので多く、これは動画投稿サイトの傾向を示している。次に、人気を表す再生回数と、他の頻度との相関関係を算出した。結果を下記の表に示す。動画の長さは人気に関係が無い（長時間の動画でも再生が多い場合がある）ことがわかった。

表1：再生回数との相関係数

	コメント数	マイリスト登録数	動画長
相関係数	0.73	0.88	0.01

### 3. 視聴者コメントに基づく検索手法の提案

ニコニコ動画では動画へのタグ付与数が10個に制限されており、そのためかタグを動画の説明やジャンルに用いることが多い。ニコニコ動画では、視聴者は面白いと思う動画に、自由にコメントを付与することができる。これは視聴者による動画の面白さを表していると考えられる。

ニコニコ動画でも、コメント数を使ったランキングが用意されており、コメントの多い順、コメントが新しい順などで検索結果を並び替えることができる。

しかし、多い順や新しい順といった単純なランキングでは、大人数が面白いと判定するものしか上位にでない。中人数の視聴者が大変面白いと太鼓判を押す動画はランキングの中位に埋もれてしまう。そこで、各コメント者の面白さ度合いと、コメント付与者の人数を考慮したランキング尺度を検討する。具体的には下記の評価尺度を検討する。

$$W1 = \frac{1}{n} \sum_{i=1..n} \log(w_i + 1) \quad W2 = \frac{1}{\log(n+1)} \sum_{i=1..n} \log(w_i + 1)$$

$$W3 = g(v) W2$$

ここで、 $w_i$  はコメント投稿者  $i$  が付与したコメントの重み（面白さ）である。 $W3$  の  $g(v)$  は、動画  $v$  の再生回数に基づく正規分布値で、 $W3$  では人気の中位の動画をより強くランキングの上位へ押し上げることを考えている。

上記の提案による面白さの評価尺度以外に、利用者が好に合致するジャンルの動画を提示する検索手法は定義していない。我々は以前からソーシャルブックマークにおけるジャンル分け手法[3]や、医学文献における関連文献提示手法[4]を提案してきた。これらはタグや文献内の出現単語の出現頻度と共起関係に基づいてジャンルを特定するものである。これらの手法を援用することで、動画のジャンル分けを行うことが可能である。

### 4. おわりに

本論文ではニコニコ動画のコンテンツを対象に、動画のメタデータ解析を行った。また、利用者コメント情報に基づく動画の検索ランキング手法を検討した。

今後は、まず提案手法の定性的あるいは定量的な評価を行いたい。次に音楽以外の動画にも提案手法を適用し、それにより全分野で適用可能であることを示したい。また、可能であれば、YouTubeの動画についても適用可能であるかを検討したい。さらに、より良い推薦手法を検討し、より良い動画閲覧を実現したい。

### 参考文献

- [1] ニコニコ動画, <http://nicovideo.jp/>, 2007.
- [2] I. Popescu, On a Zipf's Law Extension to Impact Factors, *Glottometrics* 6, pp. 83-93, 2003.
- [3] 百田信, 伊東栄典: ソーシャルブックマークに基づく情報発見, 電子情報通信学会 第19回データ工学ワークショップ (DEWS 2008), 11-15, Mar., 2008.
- [4] 園田亮, 伊東栄典: 専門用語の関係に着目した医療文献検索手法に関する考察, 情報処理学会 火の国情報シンポジウム 2009, C-4-3, Mar., 2009.