

# 楽曲に特徴的なコメントに基づく 作業用BGMからのインデックス作成

Index Generation from BGM by Using the Distinctive Comments Posted on the Songs

福田 寛公      松本 和幸      吉田 稔      北 研二  
Hirokimi Fukuda      Kazuyuki Matsumoto      Minoru Yoshida      Kenji Kita

徳島大学工学部知能情報工学科  
Faculty of Engineering, Tokushima University

The purpose of this study is to generate indexes to the BGMs based on the distinctive comments annotated to the BGM videos posted on Nico Nico Videos. Our proposed method detects the end/start positions of the BGMs considering the increase/decrease of the number of the comment data, the comments distinctive to the songs and the exclamatory expressions annotated to the videos. The evaluation result showed that the proposed method could generate correct indexes within 10 seconds error for the BGM videos with length of less than 40 minutes and with annotation of over 30,000 comments.

## 1. はじめに

ニコニコ動画や YouTube といった動画共有サイトの利用者は年々増加し、若い世代を中心とし何千万人にも及び、現在ニコニコ動画 [1] には 3000 万件を超える動画がアップロードされている。なかでも作業用 BGM の動画は動画の数も視聴者の数も多くなっている。作業用 BGM とは複数の曲が 1 つの動画にまとめられているものを指す。

しかし、アップロードされている動画の大半は内容が一目でわかるものとなっていない。視聴者が動画を実際に再生し内容を確認する作業は膨大な手間と時間がかかり快適な動画視聴を妨げている。そのため自動で動画のインデックスを作成するシステムが必要となる。

本研究では、動画に対して投稿されたコメントデータを解析することで曲名とその開始位置を示すインデックスを自動付与する手法を提案する。

## 2. 関連研究

### 2.1 ニコニコ動画における映像要約とサビ検出の試み

青木ら [2] は、映像に伴う音楽のサビ部分の検出や映像要約の生成を、コメント頻度に基づき行う手法を提案している。

ニコニコ動画でコメント数が多くなるのは「面白いシーン」であることが多く、映像要約に取り入れる価値は高いが、その面白さを理解するために必要なコンテキストが他にある場合には、十分に面白さが伝わらないことがあるとしている。また、料理番組のように調理プロセスの部分より、合間にトークが含まれている場合は、むしろそのトークが優先的に選択されてしまつて調理プロセスが捨象されてしまうことを予想している。

関連研究では、コメント頻度のみを用いているためコメント乞食の動画に対しては対応できないという問題がある。本研究ではコメント内容も考慮することによりこの問題を解決する。

## 3. 提案手法

### 3.1 特徴コメント

ニコニコ動画ではシーンに関連付けてコメントが投稿できるという特徴を生かし、曲が切り替わるごとにその曲に対応した特徴的なコメントが投稿されている。本研究ではこのコメントのことを特徴コメントと呼ぶこととする。

ニコニコ大百科 [3] など特徴コメントをまとめているサイトが存在し、多くのユーザはそれらのサイトからコピーアンドペーストによりコメントを投稿しているため、どの動画においても同じ曲ならば同じコメントが投稿される。そのため、こうした特徴コメントを検出することにより、動画中において楽曲の有無を判定することが可能であると考えられる。

表 1: 特徴コメントの一例

曲名	特徴コメント
Daydream cafe	心がぴよんぴよんするんじゃ
true my heart	きしめん
嘆きの樹	まさかの 3000 コンボイ

### 3.2 動画の分割

楽曲の開始位置付近では動画の内容が盛り上がるためコメント数が増加する傾向にある。このとき増加しているコメントの大半には感嘆表現が用いられている。感嘆表現とは、「うおおお」や「きたあああ」といった表現とする。

まずコメントデータを投稿シーンの時間軸順にソートする。コメントデータの投稿シーン時間は 100 分の 1 秒まで記録されているため  $N$  秒で丸め、 $N$  秒ごとのコメント数を見ていく。0 秒を 1 曲目の始まりとし、 $(t-1)N$  秒から  $tN$  秒 ( $t$  は正の整数) までのコメント数の変化を調べる。式 3.1 に従い、コメント数が増加している箇所を検出する。

$$x = \frac{tN \text{ 秒から } (t+1)N \text{ 秒までのコメント数}}{(t-1)N \text{ 秒から } tN \text{ 秒までのコメント数}} \quad (3.1)$$

本研究では経験的に  $x$  が 2.5 以上ならコメント数増加位置とする。

次に前述の検出した箇所についてコメントを見ていく。感嘆表現リストとのマッチングを行い、式 3.2 に従いマッチ率が低

い箇所を除外する。

$$y = \frac{\text{コメント数増加位置での感嘆コメント数}}{\text{コメント数増加位置でのコメント数}} \quad (3.2)$$

本研究では経験的に  $y$  が 0.4 以下なら除外対象とした。最終的に残った箇所を曲の切れ目と判定し動画を分割する。

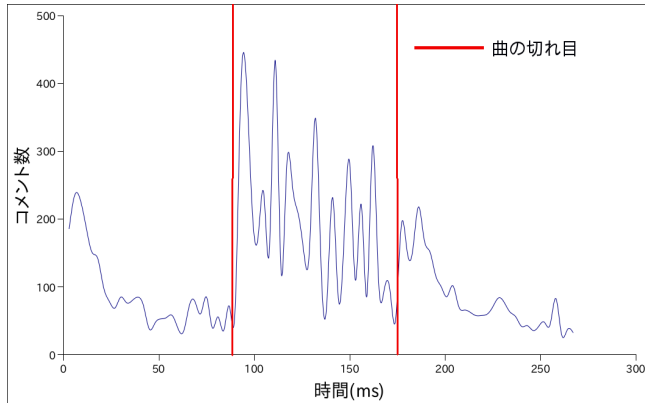


図 1: 曲の切れ目でコメント数が増加する例

## 4. 実験と考察

### 4.1 実験概要

本実験で扱う動画は、ニコニコ動画にアップロードされている動画のうち「作業用 BGM」のタグが付いているものとする。その中で動画の内容とは関係のないコメント投稿が多数を占める「数字が流れる動画」のタグが付いているもの、コメント数が再生数より多くなっているものは対象外とする。

表 2: 実験に使用したデータ

コメント数	実験に用いた動画数	動画時間の平均
10 万件以上	2	42:31
5~9 万件	3	42:29
4 万件	3	34:35
3 万件	3	24:19
2 万件	3	29:23
1 万件	3	28:23
5000 件	3	25:33
2525 件	3	46:22

実験は動画のコメント数ごとに行い、実際に動画を見て確認することにより評価を行う。

### 4.2 実験結果

表 3: 実験結果まとめ

コメント数	最大誤差 (秒)	曲名判定 (成功率%)
10 万以上	+14	23/45 (51.11)
5~9 万	+8	49/67 (73.13)
4 万	+18	55/71 (77.46)
3 万	+18	74/112 (66.07)
2 万	+2 分 24	71/108 (65.74)
1 万	+1 分 41	28/73 (38.35)
5000	+2 分 01	22/65 (33.84)
2525	+8 分 04	17/88 (19.31)

### 4.3 考察

#### 4.3.1 曲の切れ目の推定

共通して見られた傾向として曲の切れ目の判定が正解位置よりも後方にずれてしまっている。このような誤差が出る原因

として、ニコニコ動画のコメントの性質上、コメントが投稿されるタイミングは動画の内容に対し後追いになっていることが挙げられる。ユーザは動画を見てからコメントを投稿するため必ず数秒程度の誤差が出てしまう。

また動画の前半にはコメントが集中しているため誤差が小さくなっているが動画の後半になるにつれ誤差が大きくなっている。これは動画時間が長くなればなるほど顕著に見られた。しかし、ランキング形式の動画などは動画の後半になるにつれて精度が上がっている。これはランキング上位には有名な曲が出現し、コメント投稿が増加して盛り上がるため精度が上がっていると考えられる。

コメント数が 1 万件以上あり、サビメドレーのように 1 曲の時間が 10 秒前後の短い動画でなければ曲の切れ目の推定はうまく行えている。

#### 4.3.2 曲名の推定

曲名の推定についても、コメント数が 1 万件以上あり、サビメドレーでなければ高い精度で行うことが出来た。本研究で収集した特徴コメントは主に 2010 年以降のアニメ、ゲームに絞っているためそれより前に投稿された動画に対しては精度を出すことができなかった。コメント数が多い動画は投稿から年月が経っているものも多いためそのような動画に対しては、推定がうまくいかなかった。

動画に存在しない曲が出現したのは、他の曲の特徴コメントが多数投稿されていたことが原因であると考えられる。またコメント数が少ない動画では、特徴コメントは出現しているが頻度が低いいため曲名を推定することが出来なかった。しかし、コメント数が少ない動画に対し曲の切れ目は考慮せずに曲が動画の中に存在するかを実験してみた結果約 8 割の曲名を正しく推定することに成功した。

## 5. おわりに

本研究では、動画のコメント情報を用いて動画のインデックスを作成する方法を提案した。提案手法により、映像や音声情報を用いることなくコメント情報のみからインデックスを作成することが可能になった。具体的には、動画時間を 40 分以内、コメント数を 3 万件以上の条件で誤差 10 秒以内でインデックスを作成することができた。

今回の実験ではコメント数が少ない動画では曲の切れ目が推定できなかった。コメント数が少ないと特定の曲でしかコメントが増加しないことが原因であった。これを、コメント投稿シーン時間、コメント数の推移以外の素性を用いることで解決できるか検討したい。

## 6. 謝辞

本研究は JSPS 科研費 15K00425,15K00309,15K16077 の助成を受けたものである。

## 参考文献

- [1] ニコニコ動画: <http://www.nicovideo.jp/>
- [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, 2008-HCI-128/2008-MUS-75, Vol.2008, No.50, pp.37-42, 2008
- [3] ニコニコ大百科: <http://dic.nicovideo.jp/?header>