

Blog 記事からの映像コンテンツメタデータ抽出

山崎 智弘[†] 筒井 秀樹[†]
浦田 耕二[†] 福井 美佳[†]

The extraction of the metadata of video contents from blog articles

TOMOHIRO YAMASAKI,[†] HIDEKI TSUTSUI,[†] KOJI URATA[†]
and MIKA FUKUI[†]

1. はじめに

双方向放送が登場して以降、ユーザが映像コンテンツを視聴しているときにシーンに関連した話題や商品などを即座に提示する推薦サービスの実現に向けた取り組みが行なわれている。ユーザの嗜好に応じた推薦を行なうためにはコンテンツのメタデータが不可欠であるが、TV 放送のようなリアルタイム性の強いコンテンツのみならず DVD のようなコンテンツであっても、制作者がメタデータを付加することは稀である。そのため映像情報や字幕情報を解析して自動的にメタデータを獲得する研究が行なわれている。

一方今や Blog の国内開設数は 800 万件を超え¹⁾、個人からの情報発信は一般的となった。中でも映画などの映像コンテンツに対するレビュー記事の増加は個人のコンテンツ視聴スタイルを変えつつある。たとえば web 上の他者のコメントやランキングを参考にして視聴するコンテンツを選択したり、好みの DVD を何度も視聴してコメントを web 上で公開したりといったことが行なわれている。

そこで我々は、執筆されたコメントにコンテンツのシーン情報が埋め込まれていればそのコメントを分析することでコンテンツのメタデータを自動的に獲得することができると考え、映像コンテンツと Blog システムを融合したアプリケーションである“コミュテツ (CommuteTs)”を開発した。コミュテツは映像コンテンツを介して他者とのコミュニケーションを楽しむこと、またコンテンツを介したコミュニケーショ

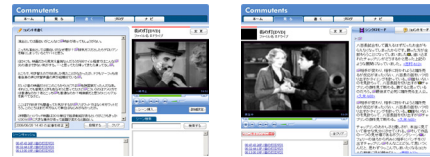


図 1 書くモード (左) と見るモード (右)
Fig. 1 Writer mode (left) and Viewer mode (right)

ンを促進することでユーザの自然な視聴スタイルからコンテンツのメタデータを獲得することを目的とする。本稿ではコミュテツの説明を行なうとともに、本システムを実際に利用して作成された Blog 記事を用いて行なったコンテンツのメタデータをシーンごとに獲得する実験の結果について述べる。

2. コミュテツ

コミュテツのようにコンテンツを介して他者とのコミュニケーションを楽しむことを目的とするシステムとしては SceneNAVI³⁾ や Synvie²⁾ などが知られているが、本システムはコンテンツの共有は行なわず、シーンを特定する情報が埋め込まれた Blog 記事のみを共有することを特徴とする。DVD の場合は VOB データの構成情報を、録画コンテンツの場合は地域・チャンネル・放送日時と映像のカット点情報をローカルで取得しサーバに問い合わせることで、コンテンツの共有は行なわずにシーン情報の同定同期を行なう。機能としては、コンテンツへの感想を Blog 記事として執筆する「書くモード」、Blog 記事を閲覧する「ブログモード」、他者のコメントを表示しながらコンテンツを視聴する「見るモード」から構成される。

図 1 左に示す「書くモード」では執筆者は SceneNAVI や Synvie のようにシーンに対する短いコメン

[†] 株式会社東芝研究開発センター知識メディアラボラトリー
Toshiba Corporation, Corporate Research and Development Center, Knowledge Media Laboratory

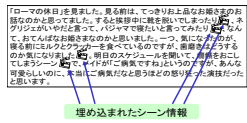


図 2 コメントの例
Fig. 2 An example of comments

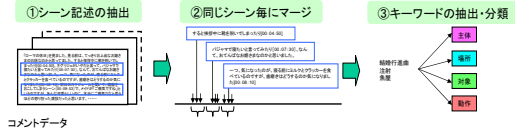


図 3 メタデータ抽出処理の流れ
Fig. 3 Process flow of metadata extraction

トを記述していくのではなく、コンテンツへの感想を通常の Blog 記事として執筆しながら任意のシーン情報を埋め込んでいくことができる (図 2)。執筆した記事はシーン情報がエンコードされた後外部の Blog サービスへ投稿される。逆に言えば外部の Blog サービスへ投稿されるので、閲覧者は本システムを利用しなくても本システムを利用して執筆された記事を読覧することが可能である。ただし本システムを利用すれば執筆者と同じコンテンツを持っている場合シーンアイコンをクリックして直接再生させることが可能となるので正確なシーンの特定が容易となる。

一方図 1 右に示す「見るモード」ではユーザが設定しておいた RSS から記事を収集し、コンテンツの再生と同期させて表示する。ただし全文では量が多く読みづらいため、記事に埋め込まれたシーン情報の前後の適切な範囲の文字列を切り出している。切り出された文字列でもシーンアイコンをクリックして別のシーンを直接再生させることができるため、記事内で関連するシーン同士を容易に比較再生することができる。

3. メタデータ獲得実験

本節ではコンテンツのメタデータをシーンごとに獲得する実験の結果について説明する。表 1 は本システムを実際に利用した 10 人のユーザによって作成された Blog 記事のコンテンツごとの記事数、埋め込まれたシーン数、文字数の例である。14 コンテンツに対して 100 記事、5945 シーンが収集された。

前述したように本システムではコメントにシーン情報を埋め込んでいくため、シーンごとのメタデータを

表 1 コメント収集結果
Table 1 Collected comments

コンテンツ	記事数	シーン数	文字数
ローマの休日	7	477	33587
パッチギ!	10	631	54644
...
総数	100	5945	492032

抽出するためにはコメントからそれぞれのシーンに関連した箇所を自動的に切り出す必要がある。メタデータ抽出処理の流れを図 3 に示す。今回の実験では文単位で切り出しを行ない、前後の文を含めるか含めないかは段落や接続詞の情報を用いて決定した。その後切り出されたシーン記述のうち時刻の差が閾値以下など同じシーンに言及していると考えられるものをひとまとめにし、形態素解析を用いてそれぞれのシーンごとに頻出するキーワードを抽出し、その後抽出されたキーワードが主体、場所、行動などのどれに相当するかを辞書を用いて自動的に分類した。表 2 はローマの休日に関連するコメントからキーワードとシーンの説明表現をシーンごとに抽出した結果の例である。アン王女など字幕にあらわれるキーワードだけでなく魚屋、ヴェネツィア広場、暴走など字幕にはあらわれないキーワードも正しく抽出されていることが確かめられた。これはユーザの自然な視聴スタイルからコンテンツのシーンごとのメタデータが実際に獲得できたことを示している。

4. おわりに

本稿では映像コンテンツと Blog システムを融合したコミュニティを提案した。また本システムによる Blog 記事から実際にコンテンツのメタデータをシーンごとに獲得することができることを示した。今後はメタデータ抽出の精度をシーン検索などの応用を通じて検証・向上していく予定である。

参考文献

- 1) ブログ及び SNS の登録者数 (平成 18 年 3 月末現在). <http://internet.watch.impress.co.jp/cda/news/2006/04/13/11636.html>
- 2) 山本大介ほか: Synvie: ブログの仕組みを利用したマルチメディアコンテンツ配信, 情報処理学会第 68 回全国大会, 2006.
- 3) 東正造ほか: 多種類のテレビ映像を対象とした映像シーン運動型掲示板におけるコミュニケーションの分析, グループウェアとネットワークサービスワークショップ 2006, pp.31-36.

表 2 「ローマの休日」
Table 2 "Roman Holiday"

00:56:27	主体	アン王女, おばさん, ブラドリー
	場所	魚屋, 街中, 市場
	行動	買い物, 探索, カット
	説明	王女が街中を冒険するシーン
01:15:37	主体	アン王女, ジョー, 警察
	場所	街中, ヴェネツィア広場
	行動	暴走, 連行
	説明	説明不要なほど有名な街中でスクーターを乗り回すシーン