

## 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システム

山本 大介†      清水 敏之†      大平 茂輝†      長尾 確†

†名古屋大学 情報科学研究科

〒464-8603 名古屋市千種区不老町 IB 電子情報館南棟 395 号室

E-mail: yamamoto@nagao.nuie.nagoya-u.ac.jp

あらまし 本研究では、マルチメディアコンテンツ配信にブログの仕組みを取り入れることによって、マルチメディアコンテンツを話題とした コミュニケーションを支援するシステムを開発した。さらに、コンテンツ閲覧者による自然な知的活動(コンテンツに関するレビュー記事を書くなど)から、コンテンツに関する知識を RDF 形式のアノテーションとして獲得・配信する仕組みを提案する。

キーワード ブログ, アノテーション, コミュニティ支援

## Synvie: A Weblog-style Multimedia Contents Distribution System.

Daisuke Yamamoto†      Toshiyuki Shimizu†      Shigeki Ohira†      Katashi Nagao†

†Graduate School of Information Science, Nagoya University

Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan

E-mail: yamamoto@nagao.nuie.nagoya-u.ac.jp

**Abstract** In this research, we propose the multimedia contents distribution system that supports user's communication concerning multimedia contents by applying the mechanism of Weblog. In addition, we propose the mechanism that acquires knowledge of contents as annotation of the RDF form from user's natural intelligent activity, such as editing a review article.

**Key words** Weblog, Annotation, Community support

### 1 はじめに

近年、インターネットの発達と共に、映像・音楽などのマルチメディアコンテンツが Web 上で頻繁に配信されている。それに伴い、コンテンツをいかに効率よく配信するかといった問題が顕在化してきている。その一方で、ブログや SNS, Wiki などの登場により個人や Web コミュニティからの情報発信が一般化し、その影響力も増している。そこで本研究では、マルチメディアコンテンツとそれらを取り巻く Web コミュニティとを効果的に融合させ、それらのコミュニティにおけるユーザの自然な知的活動からコンテンツに関する知識をアノテーション [4]

として獲得・利用することを目的としている。具体的には、ユーザが一般的な Web ブラウザを用いてコンテンツを閲覧し、容易にコンテンツの内容に対する感想や評価などの情報の関連付けを支援する仕組みを作成することによって、閲覧者らによるコミュニケーションを支援する。さらには、コンテンツの内容とこれらのコミュニケーションとを詳細に結びつけることによって、コンテンツに付随する様々な情報をアノテーションとして獲得する。これらのコミュニケーションは単体のコンテンツのみに閉じているのではなく、ブログの仕組みを利用することによって、Web 全体を対象とするより広がりをもつコ

コミュニティの構築を支援する。これらを実現するための具体的なプラットフォームとして、Synvie というシステムを開発した。なお、本論文では主に映像コンテンツを対象として述べているが、同様な考え方は音楽や画像など他のマルチメディアコンテンツに対しても有効である。

## 2 コンテンツの Weblog 化

本研究では、既存の Weblog エントリーと同様な形でマルチメディアコンテンツを配信するための枠組みを提案する。Weblog で映像コンテンツを配信することを目的としたサイトは既にいくつか存在しており、一般にビデオブログと呼ばれている。現状のビデオブログは、映像の任意の時間区間に対してのコメントが付与できないなど、いくつかの課題や問題点が指摘 [5] されており、必ずしも十分に機能を果たしているとはいえない。

### 2.1 本研究における Weblog の定義

Weblog は、インターネット上で頻繁に更新される、主にテキストと画像からなる HTML 形式で記述された記事を効率よく配信するための仕組みである。Weblog の詳細な定義については諸説あるが、本論文では技術的側面から見て以下の 4 つの要件（ここでは Weblog 化要件と呼ぶ）を満たす Weblog を対象とする。

1. Permalink[1] : エントリーを一意かつ永続的な URI により特定する仕組み。
2. Annotation : エントリーに対して閲覧者が感想やコメントなどを付与できる仕組み。
3. Trackback[2] : ユーザがエントリー B を参照してエントリー A を記述する際、エントリー B からエントリー A へのリンクを明示的に生成する仕組み。これにより、エントリー間での双方向リンクの形成が可能になる。
4. XML Feeds : RSS 1.0<sup>1</sup>や Atom<sup>2</sup>など、更新情報を XML 形式で配信する仕組み。

これらの仕組みを用いて、各々のエントリーを Trackback などのリンクで結びつけることにより、Weblog エントリーを中心としたコミュニティを活性化することができる。なお、これらの仕組み

<sup>1</sup><http://web.resource.org/rss/1.0/>

<sup>2</sup><http://www.ietf.org/html.charters/atompub-charter.html>

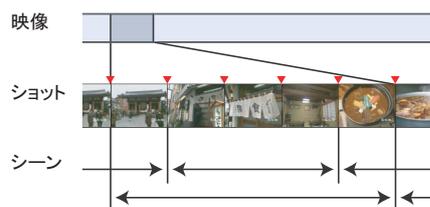


図 1: 映像のモデル。映像は複数の連続するショットからなり、シーンは複数のショットからなる。

は、MovableType(<http://www.movabletype.org/>) や Blogger(<http://www.blogger.com/>) などの主要な Weblog ツールの多くで既に実現・規格化されており、異なる Weblog ツール間での相互利用が可能である。

つまり、これらの仕組みをマルチメディアコンテンツに対して効果的に適用することによって、既存の Weblog ネットワークにマルチメディアコンテンツを参加させることができ、コンテンツを中心としたコミュニティを活性化させることが期待できる。

### 2.2 映像のモデル化

本研究では、図 1 のように、映像は複数のショットからなるリストであると定義する。ショットは、一般に映像のカット（切れ目）から次のカットまでの時間範囲を示すが、必ずしもカットが意味的な内容の切れ目であるとは限らないので、一定時間間隔に分割した単位を仮想的なショットとして仮定しても良いこととする。また映像を Web 上でより扱い易くするために、それぞれのショットの内容を表すサムネイル画像をあらかじめ用意する。シーンとは、複数の連続するショットからなり、意味的につながりを持っているものと定義する。ひとつのショットが複数のシーンに属することも許す。

### 2.3 映像コンテンツとその内容に対する Permalink

Weblog 化のための一つ目の要件として、映像コンテンツ全体や任意のシーンに対して Permalink を記述可能である必要がある。映像コンテンツ全体に対する Permalink は以下のように、固有の ID を用いた URI を記述する。

`http://[server]/[content ID]/`

また、任意のシーンに対する Permalink は、以下のように固有の ID とその時間区間を記述する。複

数の時間区間に対する Permalink を記述する場合は、コマンドで区切って複数記述すればよい。RDF[3]の主語として Permalink を用いる場合には一意である必要があるため、begin time の早いほうから順に並べ替えて記述する。

```
http://[server]/[content ID]/
  [begin time]-[end time],
  [begin time]-[end time], ...
```

これらの仕組みにより、映像の任意の時間区間に対して、固有の Permalink を記述することができる。

## 2.4 映像コンテンツとその内容に対するアノテーション

Weblog 化のための二つ目の要件として、ユーザがコンテンツの任意のシーンに対して容易にコメントの付与などのアノテーションを可能にする仕組みが必要である。そのために、筆者らが以前の研究で作成したオンラインビデオアノテーションシステム iVAS[6] の仕組みを発展させて利用する。

ユーザは、ネットワークからアクセス可能な任意の映像コンテンツに対して、Web ブラウザを用いてアノテーション及び閲覧を行う。本研究では、シーンに対してコメント情報を記述することをテキストアノテーションと呼び、二種類のインタフェースを提案する。一つ目は、図 2 に示すように、映像の任意のショットに対してコメントを付与できる簡便なインタフェースであり、映像の閲覧を継続したままアノテーションを付与可能である。しかしながら、このインタフェースでは映像中の任意の領域を対象としたアノテーションが付与できない。二つ目は、図 3 のような、矩形範囲に対してコメントを付与するためのインタフェースである。対象となるシーンの静止画像に対して、マウスで矩形範囲を選択した後にコメントを付与する。これにより、映像の任意のショットの矩形領域を対象としたアノテーションの付与が可能になる。このインタフェースでは、映像の閲覧を一時的に停止する代わりに、より詳細なアノテーションを付与可能である。また、より簡易なアノテーションとして、筆者らは iVAS[6] において、印象アノテーションというものを提案した。印象アノテーションでは、映像の任意の時間に対してマウスクリックであらかじめ用意された閲覧者の主観的な印象を表すボタンを押すことによって統計的に評価する仕組みである。



図 2: 映像の任意のショットに対するアノテーションおよび視聴インタフェース。ユーザはショットに対してコメントを付与可能である。また、現在の映像に同期したアノテーションを表示可能である。



図 3: 映像の任意のショット及び矩形範囲に対するアノテーションインタフェース。現在のショットに対応する静止画像に対してマウスで矩形範囲を選択し、コメントを投稿可能である。

ユーザはこれらの仕組みを用いることにより、映像コンテンツに対して、電子掲示板感覚で他のユーザとコミュニケーションを図ることが可能になると同時に、様々な種類のアノテーション情報を映像コンテンツに付与することが可能になる。これにより、映像コンテンツに対する多目的に利用可能なアノテーションの蓄積が可能になり、検索や要約、コミュニティ支援などの様々な応用に利用できる。

## 2.5 映像コンテンツとその内容に対する Trackback

Weblog 化のための三つ目の要件として、Trackback の仕組みを映像コンテンツに適用する必要がある。Trackback は、通常のサイト単位の相互リンクとは異なり記事単位のリンクであるため、よりピンポイントにコンテンツに依存したリンクを張ることができる。さらに、このリンクは人間の主観によ

て張られるものであり、記事内容を意識した関連性の強いリンクとみなすことができる。

Trackback の形式は規格化 [2] されており、相手側のエントリーの Trackback URI に対して、サイトの名前、エントリーのタイトル、エントリーの概略、エントリーの URI 情報を送信する。これを Trackback ping と呼ぶ。Trackback URI は通常以下のような形式となっており、ユーザはこの URI を直接入力することによって相手先エントリーを指定する。

```
http://[server]/[entry ID]
```

ビデオブログの Trackback の仕組みは、Weblog と同じであり、Trackback ping を送信することによって成立するべきである。映像コンテンツの任意のシーンに対して Trackback を送信するためには、そのコンテンツの ID、シーンの開始時刻・終了時刻の情報も送信する必要がある。Trackback URI 以外は厳密に規格化されているため、我々は Trackback URI にこれらの情報を埋め込み、Permalink と同様に以下の形式に拡張した。

```
http://[server]/tb/[content ID]/  
[begin time]-[end time]
```

また、複数のシーンに対してトラックバックを付与したい場合には、Permalink と同様に、コンマで区切って複数の [begin time]-[end time] を記述すればよい。

### 3 レビュー記事の生成と配信

次に、既存の Weblog とビデオブログを積極的に結びつけ、映像コンテンツとそれを取り巻くコミュニティを活性化させるための仕組みを提案する。具体的には、ユーザが映像コンテンツの任意のシーンを容易に引用し、そのコンテンツに関する記事（レビュー記事）の作成を支援する仕組みを提供する。これにより、ユーザは自身の Weblog においてコンテンツの宣伝を行うことができ、コンテンツの流通の促進が期待できる。映像コンテンツの引用は、シーンを示すサムネイル画像とそのシーンに対するリンク及びコメントからなる。この形式を複数組み合わせる Weblog エントリーを記述する。

#### 3.1 レビュー記事の生成

ユーザはコンテンツを閲覧する際、2 章 2 節のアノテーションインタフェースを用いて閲覧し、自身

にとって興味のあるシーンに対しては何らかのテキストアノテーションや印象アノテーションを付与しているものとする。もしも、ユーザがレビュー記事を作成したいと考えた場合、これらの手がかりを残したショットを対象とした記事を記述したいという欲求があると考えられる。

そこで、システムは、これらのアノテーション履歴から、アノテーションを施したショットを対象としたレビュー記事テンプレートを自動生成する。レビュー記事テンプレートは、テキストアノテーションを施したショットに対してはそのショットのサムネイル画像及びコメント、印象アノテーションを施したショットに対してはそのショットのサムネイル画像と空欄テキストフィールドからなる。また、ショット及びコメントは次節で述べるインタフェースを用いて修正可能である。

#### 3.2 レビュー記事の編集

ユーザは、Weblog などで通常のエントリーを書くのと同様に、一般的な Web ブラウザを用いて、レビュー記事テンプレートの編集が可能である。

本研究では、二つの編集インタフェースを提案する。一つ目は、シーン伸縮型の映像レビュー記事の編集インタフェース (図 4) である。これは、対象シーンを時間的に前後にショット単位で伸縮させることに対象シーンの修正・変更が可能であり、より正確に選択することが可能なインタフェースである。このインタフェースを用いて選択された連続する複数のショットはひとつの仮想的なシーンとして捉えることが可能であるため、シーンの流れやストーリーを対象とした記事を記述するのに適したインタフェースである。

二つ目は、ドラッグアンドドロップ型の映像レビュー記事の記述インタフェース (図 5) である。過去にユーザが施したテキストアノテーションや印象アノテーションに対応するショットが右側のストックに保持されており、その中から任意のショットをドラッグアンドドロップ形式で複数選択し、その複数のショットに対してコメントを付与することが可能なインタフェースである。これは、複数の連続しないショットに対してコメントを記述することに適したインタフェースであり、シーンやストーリーよりも特定のオブジェクト（たとえば特定の人物など）を対象とした記事を記述するのに適したインタフェースである。



図 4: シーン伸縮型の映像レビュー記事の編集インタフェース。シーン伸縮ボタンを押して対象シーンを時間的に前後に伸縮させることによって、正確に対象シーンを提示・選択可能であり、対応するコメントの編集も可能である。



図 5: ドラッグアンドドロップ型の映像レビュー記事の編集インタフェース。過去に投稿したアノテーション履歴から、任意の複数のショットを選択し、対応するコメントの編集も可能である。

ユーザはこの二つのインタフェースを使い分けながら記事を作成可能である。

編集されたレビュー記事は、既存のブログへ XML-RPC や Atom API の仕組みを用いて投稿する。

### 3.3 レビュー記事からアノテーションの獲得

レビュー記事の編集結果には様々な知識が含まれており、二つの観点からアノテーションとして知識の獲得が可能である。

一つ目は、アノテーションのコメント情報について考察する。映像を閲覧しつつリアルタイムに投稿されたテキストアノテーションには、誤字脱字や、言葉足らずなコメントが少なからず存在する。レビュー記事を執筆することによって、これらのテキ

スト情報を見直し、テキストの修正や補完が行われていることが期待できる。また、印象アノテーションには、本来テキスト情報は含まれていないが、ユーザに印象アノテーションを施したシーンに対してコメントの付与を促すことによって、テキスト情報が付与される可能性もある。つまり、映像のシーンに対して、より質が高くより量が多いテキスト情報の獲得が期待できる。

二つ目は、アノテーションの対象シーン情報について考察する。誤って選択された対象シーンの修正を期待できるだけでなく、映像の構造情報の取得が可能になる。シーン伸縮型インタフェースによって選択された連続するショットからなる対象シーンでは、それに対応するコメント内容という観点に基づいた一つの仮想シーンとみなすことができる。また、ドラッグアンドドロップ型インタフェースを用いて選択されたショットの集合は、対応するコメントの意味内容という観点に基づいて、関連性があると考えられる。複数のレビュー記事を統合的に解析することによって、ショットの連続性と、ショットの関連性という二つのパラメータを用いた、映像の隠れた構造を発見することが可能になり、映像の構造化に関する知識を獲得することができる。また、これらの構造はグラフ構造で表現される。

## 4 アノテーションの記述と流通

本システムでは、コンテンツに付随する様々なメタデータを積極的に配信する。XML Feed に代表されるコンテンツの更新履歴やサマリー情報といった情報をただ単に配信するだけではなく、コンテンツに付随するアノテーション及び閲覧者に関する情報も積極的に配信する。コンテンツに付随するメタ情報はコンテンツの内容に依存した情報である場合が多く、これらの情報を解析することによって間接的にコンテンツの内容を解析することが可能になる。

### 4.1 アノテーションの配信

映像コンテンツを Weblog 化するための四つ目の要件として、XML 形式による情報の配信する仕組みを提案する。配信される情報は、コンテンツの更新情報だけでなく、アノテーション情報も配信する必要がある。アノテーションは 3 章 3 節で述べたようにグラフ構造で記述することが望ましいので、グラフ構造の記述に適した RDF 形式により記述する。具体的には、rdf:about 属性に一意的映像コンテン

ツの URI とアノテーションの ID を指定することによって、以下のような形式で表現する。

```
<annotation rdf:about="http://[server]/
    [content ID]#[annotation ID]">
  <target rdf:resource="http://[server]/
    [content ID]/[begin time]-[end time]">
  <annotator id="maya@synvie.net"/>
  <dc:date>2005-10-22T21:03:52</dc:date>
  <description>コメント</description>
  ... Other Annotation XML Data ...
</annotation>
```

target 要素の rdf:resource 属性には、先に述べた任意の映像コンテンツを指す Permalink を記述する。また、テキストアノテーションの場合はコメントの内容を表す description 要素や、アノテータ情報を記述する annotator 要素などを定義することによって記述可能である。これらは一意的な URI や ID が付与されているため、様々なビデオブログサーバに分散する「誰」が「いつ」「どのコンテンツ」の「どのシーン」に「どのような内容」を付与したかというアノテーションを機械が可読な形で配信・収集することが可能になる。

## 4.2 Weblog ネットワーク

Weblog の仕組みを適用することによって、マルチメディアコンテンツを図 6 のような Weblog ネットワークに参加させることが可能になる。これにより、コンテンツ・人・Weblog などが互いにリンクした詳細かつ広範囲な関係をグラフ構造により構築することが可能になる。これにより、コンテンツに付随する様々な知識を発見するためのフレームワークとして機能することが期待できる。

## 5 おわりに

本論文では、Weblog の仕組みを映像コンテンツの任意のシーンに対して適用する仕組みを提案した。これにより、映像コンテンツを Web 上で効率よく扱う枠組みを提案しただけでなく、より詳細かつ多くのアノテーションを獲得することが可能になった。これらのアノテーションは、二つの観点により映像を構造化する。一つは、レビュー記事を執筆することによってそれぞれのショット間の意味的な関係の抽出が期待できる。もう一つは、Weblog の仕組みを用いることによって Web 上の任意のコンテンツ

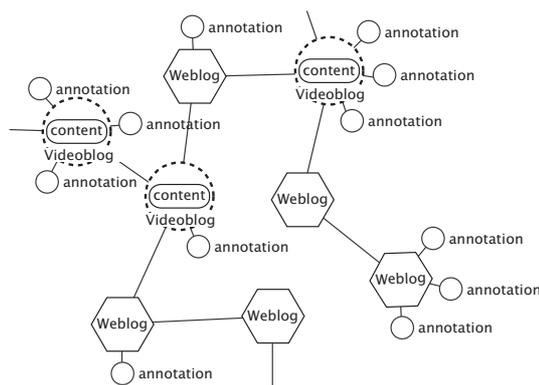


図 6: Weblog ネットワーク。

との意味的な関係の抽出が期待できる。これらはグラフ構造の記述が可能な RDF 形式で記述・配信・収集することによって、映像コンテンツに対するアノテーションの RDF データベースを構築することが可能になる。

なお本システムは、2005 年度前期の IPA 未踏ソフトウェア創造事業による支援を受け、現在開発中 (<http://synvie.net/>) である。今後の課題として、システムを公開し、より実践的なアノテーションの収集・蓄積とそれに基づく応用に関する研究を行っていく。

## 文 献

- [1] Aimeur, E., Brassard, G. and Paquet, S.: Using Personal Knowledge Publishing to Facilitate Sharing Across Communities, *Proceedings of the Twelfth International World Wide Web Conference (WWW2003)* (2003).
- [2] Benjamin and Trott, M.: *mttrackback - TrackBack Technical Specification*, movabletype.org, <http://www.movabletype.org/docs/mttrackback.html> (2002).
- [3] Manola, F. and Miller, E.: *RDF Primer, W3C Recommendation 10 February 2004*, W3C, <http://www.w3.org/TR/rdf-primer/> (2004).
- [4] Nagao, K., Shirai, Y. and Squire, K.: Semantic Annotation and Transcoding: Making Web Content More Accessible, *IEEE MultiMedia*, Vol. 8, No. 2, pp. 69–81 (2001).
- [5] Parker, C. and Pfeiffer, S.: Video Blogging: Content to the Max, *IEEE Multimedia*, Vol. 12, No. 2, pp. 4–8 (2005).
- [6] 山本大介, 長尾 確: 閲覧者によるオンラインビデオコンテンツへのアノテーションとその応用, *人工知能学会論文誌*, Vol. 20, No. 1, pp. 67–75 (2005).