映像を話題としたコミュニティ活動支援に基づくアノテーションシステム

山本大介[†] 増田智樹[†]
大平茂輝[††] 長尾確[†††]

本論文では，映像コンテンツに関連したブログコミュニティから映像コンテンツに関する意味情報をアノテーションとして獲得する仕組みを提案する．具体的には，ユーザが手軽に任意の映像シーンを引用したブログエントリーの作成が可能なツールを開発した．その編集履歴から映像シーンとブログエントリーの文書構造とを関連付けることにより，アノテーションを獲得する．これらを実現する具体的なシステムとして Synvie を開発した．さらに，本システムを用いた実証実験を行い，取得されたアノテーション情報の特性を解析することによって，本システムの有用性を確認した．また，これらのアノテーションに基づく具体的なアプリケーションの例として，映像シーン検索システムを提案する．

A Video Annotation System Based on Community Activities

Daisuke Yamamoto,[†] Tomoki Masuda,[†] Shigeki Ohira[††]
and Katashi Nagao[†††]

In this paper, we propose a mechanism which acquires semantics of video contents as annotations from related Web communities. In particular, we have implemented a Web-based tool which user can easily create a Weblog entry quoting video scenes. This tool can acquire relationships which associate multiple video scenes with a document structure of a Weblog entry from editing histories. We have developed an online video quotation system "Synvie." Moreover, we have analyzed real annotation data which were accumulated using the public beta service which we are providing, and confirmed the usefulness of our system. As an example of applications based on these annotations, we present a video scene retrieval system.

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・共有されている．それらのコンテンツは専門家が作成したコンテンツだけではなく，一般ユーザが撮影・作成したコンテンツも爆発的に増加しており，それらのコンテンツをいかに効率よく配信・管理・検索するかといった問題が顕在化している．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，影響力も増している．

† 名古屋大学 情報科学研究科
Graduate School of Information Science, Nagoya University
†† 名古屋大学 エコトピア科学研究所
EcoTopia Science Institute, Nagoya University
††† 名古屋大学 情報メディア教育センター
Center for Information Media Studies, Nagoya University

映像コンテンツの内容検索や要約などの応用を実現するためには，映像シーンに対応するメタ情報（アノテーション）の取得が必要[8] である．とりわけ，映像シーンの内容に関連したキーワードの抽出や，そのシーンの重要度の推定が有効である．映像シーンに関連したアノテーションの取得に関する従来手法としては，映像認識や音声認識等の自動解析技術を利用する自動アノテーション方式[13] や，専任の作業者が専用のツールを用いてアノテーションを作成する半自動アノテーション方式[4),12] などがある．しかしながら，とりわけ個人が作成したコンテンツの場合，手ぶれ・ピンぼけ・雑音・不明瞭な声などといった撮影者の技能の問題や，カメラ付き携帯電話やデジカメといった撮影機器の性能問題から映像や音声の品質のばらつきが大きく自動解析は限定的にしか利用できない．また，専任の作業者による半自動アノテーションを行うためには，視聴者が限定され，費用対効果の問題から，全ての映像コンテンツに対するアノテーションを施すこ

とは困難である．

そこで本研究では，映像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る仕組みを提案し，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[9] として獲得・蓄積・解析することを目的としている．具体的には，二つのコミュニケーション手段を提供する．一つ目は，映像コンテンツの任意のシーンに対して，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掲示板型コミュニケーションの仕組みであり，二つ目は，任意の映像シーンを引用したブログエントリーの生成を支援するブログ型コミュニケーションの仕組みである．これらの仕組みを作成することによって，ユーザ同士の映像を題材とした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．このような方式ならば，映像の質やアノテーションコストに左右されず，上述した自動・半自動アノテーションの問題を回避できる．

そこで，本論文では，映像シーンへのアノテーションの仕組み，映像シーン単位でのコンテンツの引用に基づくブログエントリーからのアノテーション取得方法の提案，コミュニケーションに特化した具体的なインタフェースの提案，及び，それらの仕組みを実装した Synvie[19),20)] というシステムを開発した．さらに，Synvie の公開実験に基づく分析・評価を行い，コミュニケーションから得られるアノテーションを用いたアプリケーション作成のための指針を提示する．

2. 関連研究

映像に対するコメント付与やブログへの引用といったサービスは YouTube☆や Google Video☆☆など既にいくつか限定的ながら提供されている．これらのサービスでは，コンテンツ閲覧者がその映像に対してのコメント付与による掲示板型コミュニケーションや個人のブログへの埋め込みなどが日常的に行われている．これらを映像に対するアノテーションとして捉えることは可能であるが，アノテーションの対象がコンテンツ単位であるなど粒度が荒く，映像のシーン検索などの応用に利用することは困難であり，限定的な応用にしか利用できない．

映像シーンに対するアノテーションの仕組みとして、MPEG-7[7)] が良く知られている。MPEG-7 では、おもに単体の映像コンテンツに対して専任の作業者が，映像シーン検索や要約等の応用を実現するための有用で信頼性の高い情報を記述するための枠組みであり，不特定多数のユーザが Web 上で自由コメントを執筆することは想定されていない。そのため，MPEG-7 の記述を目的とした既存のツールを流用することは困難である．

映像のシーン単位に対するアノテーションの例としては，iVAS[18)]，SceneNavi[14)] などが存在するが，アノテーションとしての利用や検証が十分ではない．また，映像コンテンツとブログなどの外部の Web サイトとを詳細に関連付け，そこからアノテーションを抽出しようとする試みはない．コンテンツに関連するコミュニティは，ブログ等の他の関連するコミュニケーションシステムにも分散する可能性が高く，そこに重要な知識が存在している可能性も高い．

また，映像と外部の Web サイトを関連付けてアノテーションを抽出する研究の例としては，Dowman ら[5)] による，ニュース映像の音声認識結果と CNN の Web ニュースステキストの内容を比較することによって自動的に該当するニュース記事を特定し，その記事から映像コンテンツに関連した情報の取得を試みる仕組みがあるが，ニューストピック単位での関連付けであるため粒度が荒く，映像コンテンツはニュース記事に限定され，また，音声認識や言語解析結果に極めて依存したリンクであるため，そのリンク自体の精度や再現性も高くない．

また，画像にタグをゲーム感覚で付与する仕組みとして，Google Image Labeler☆☆☆がある．これは，対戦型のオンラインゲームであり，対戦者が互いに一つの画像に対して連想するだろうタグを入力し，一致したタグの数に応じて得点が増えるゲームである．タグ付与にかかる一人当たりの人的コストが最小化できるばかりか，エンタテインメントとしての側面も持ち合わせた仕組みである．映像を話題としたコミュニケーションもエンタテインメントの一種であると考えれば，我々が提案する Synvie においても同様な効果が期待できる．

3. アノテーションと引用のためのプラットホーム

一般に映像コンテンツはバイナリデータであるため，

☆ http://www.youtube.com/
☆☆ http://video.google.com
☆☆☆ http://images.google.com/imagelabeler/

意味内容を考慮した上で柔軟に扱うことは困難である．コンテンツを取り巻くコミュニティからアノテーションを効率よく取得するためには，機械や人間にとっても扱いやすい枠組みを提供することが望ましい．しかしながら，現状では映像コンテンツを異なるサイト間で横断的に扱うためのプラットホームが存在しない．そこで，HTML コンテンツの管理・配信・機械的処理等で一定の成果をあげているブログの仕組みを参考にして，映像コンテンツの配信とアノテーションの枠組みについて考察する．

3.1 ブログに学ぶ

ブログでは，エントリー毎に，Permalink[1]，Trackback[3] 等の仕組みを実装することによって異なるサイトにまたがるエントリー間のリンクや引用を可能にしている．また，XML Feed[2),6)] の仕組みを利用することによって，コンテンツの情報を機械が理解可能な形で積極的に配信している．さらに，エントリーに対してコメント投稿機能を用意することによってユーザからのフィードバックを取得可能である．これらの仕組みを実装することによって，口コミによって，ブログコミュニティは急激な発展を遂げることが可能になり，RSS リーダやブログ検索などの様々な応用を生み出してきた．Parker[11] によると，ブログの仕組みを映像コンテンツに適用することによって，ビデオブログ検索やビデオブログ配信などといった高度なアプリケーションが実現できると述べている．我々は，さらにこれらの仕組みを映像コンテンツのシーンに対して適用することによって，映像シーン単位でのアノテーションや引用を実現する．これにより，既存のブログと親和性が高い，高度なアプリケーションが実現できるのではないかと考えた．

3.2 映像シーンとショットの定義

本研究では，図 1 で示すように，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，長いショットは一定時間間隔に分割しても良いこととする．本システムでは間隔を 2 秒とした．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．

3.3 映像シーンに対する Permalink

映像の任意のシーンに対してアノテーションなどの処理を施すためには，それらのシーンに対して固有の Permalink を記述できる必要がある．そこで，本研究では梶ら[16] によって提案されている Element Pointer の仕組みを採用した．Element Pointer は任意のコンテンツの部分要素に対して URI を関連付ける仕組みであり，それぞれのコンテンツの URI が一意であることが保証されている．

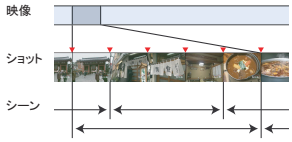

図 1 本論文における，映像のシーンとショットの定義
Fig. 1 Definition of Video Scene and Video Shot

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，コンマで区切って複数記述する．

```
http://server/[content_id]#epointer(
urn:aps:timeline(begin,end),
urn:aps:timeline(begin,end), ...)
```

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．

4. 映像シーンへのアノテーション

アノテーションには，従来からあるコンテンツの属性情報や構造情報・意味情報など，検索や要約等の応用を目的とした主次的なアノテーションの他に，副次的なアノテーションが存在すると考えている．副次的なアノテーションでは，コンテンツに付随するユーザの自発的なコミュニケーションや，コンテンツを話題としたブログエントリーの作成等のコミュニティ活動の副産物としてアノテーションの獲得を目指す．

本システムでは，様々な種類の映像コンテンツに対してアノテーションを付与することを想定している．そのため，コンテンツの種類やユーザの目的によってアノテーションインタフェースを使い分けることが有効であり，いくつかの具体的なインタフェースについて説明する．

4.1 映像シーンへのコメントアノテーション

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

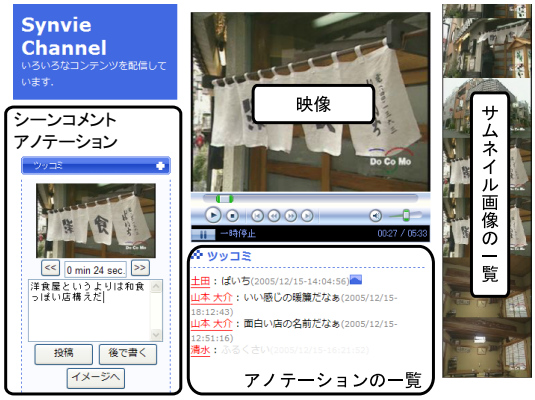

図 2 シーンコメントアノテーション．ユーザは現在再生中の映像付近の任意のショットに対してコメントを付与可能である．また，現在の映像に同期したアノテーションを表示可能である．
Fig. 2 Scene Commentary Annotation

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[18] 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の投稿及び共有を行う．本研究では，シーンに対してコメントを記述することをシーンコメントアノテーションと呼ぶ．図 2 に示すように，映像の現在再生中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．

これにより，ユーザは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，関連情報を提示したい，感動を共有したい等という欲求を満たすことが可能になる．想定するアノテーションの内容としては，映像シーンに関連した有用情報やＵＲＬ，感想などで，比較的短いコメントである．アノテーションの質はそれほど高いものを想定しておらず，このアノテーションをきっかけとした，次章で述べるシーン引用に基づくブログ執筆を促すことを考えている．

4.2 映像シーン領域へのコメントアノテーション

シーン領域コメントアノテーションとは，図 3 のように，任意の映像シーンの任意の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は，映像の閲覧を一時的に停止する代わりに，より詳細で対象が明確なアノテーションを付与可能である．

これは，映像の特定領域に対してのみコメントを記述したい時に有用なインタフェースである．想定されるアノテーションの内容としては，映像上の登場人物やオブジェクトの名称の記述，テロップの書き下し，見落としがちな部分についての注釈などが考えられる．シーンコメントアノテーションよりは説明的な記述が想定される．

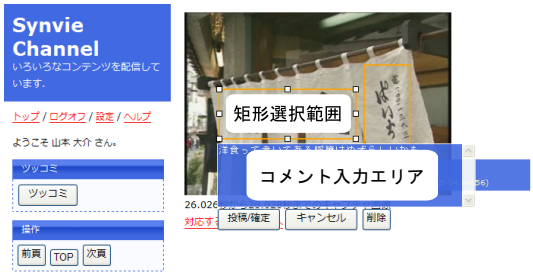

図 3 シーン領域コメントアノテーション
Fig. 3 Scene Region Commentary Annotation

4.3 映像シーンへのボタンアノテーション

次に，映像シーンに対するより簡便なアノテーションとして，二種類のボタン押下によるアノテーションを提案する．一つは，映像に対するマーキングとしての機能であり，任意のシーンに対して “チェック” を行う仕組みである．これは，次章で述べる映像シーンの引用の手がかりとして用いられ，他のユーザとの共有は行わない．二つ目は，iVAS[18] において提案されたシーンボタンアノテーションである．シーンボタンアノテーションでは，映像の任意の時間に対してマウスを用いて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．本システムでは，nice と boo の二種類のボタンを用意した．インタフェースを図 4 に示す．

本アノテーションでは，ユーザにとって興味深いシーンに対してより多くのボタンが押下されることを期待している．具体的には，面白いシーンや有用なシーン，映像的表現が面白いシーンや批判が集中しやすいシーンに対して多くのボタンが押下されると考えている．

4.4 コンテンツへのアノテーション

映像シーンに対するアノテーションだけでなく，YouTube などの従来の動画共有サイトで一般的に行われている，コンテンツ全体に対するコメント投稿の機能も実装した．これによって取得されるアノテーションを，コンテンツコメントアノテーションと呼ぶ．また，タイトル情報などあらかじめコンテンツに埋め込まれているメタデータも，コンテンツの内容を示す重要な情報でありアノテーションとして扱う．

想定されるアノテーションとしては，コンテンツ全

図 4 シーンボタンアノテーション
Fig. 4 Scene Button Annotation

体に対するコメントや感想・評価などである．

5. 映像シーンの引用に基づくアノテーション

一般的にユーザがコンテンツを閲覧し，そのコンテンツが有益で面白いと感じた場合，自身のブログ上でそのコンテンツへの URL を付与した紹介記事の執筆を行うことがしばしば見受けられる．これは，金銭的な見返りを期待しないユーザの自然で自発的な行動である．これらの記事の中にはコンテンツの内容について詳細に記述している記事も存在する．それらの記事の内容と映像コンテンツとを詳細に関連付けることができれば，コンテンツの要素に対するアノテーションとして捉えることが可能になる．Synvie では特に個人のブログエントリーへの引用を支援する仕組みを提供し，その仕組みを利用したユーザの詳細な編集履歴を蓄積することによって，ブログエントリーの文章構造と映像のシーン構造とを関連付けたアノテーションの抽出を可能にする仕組みを提案した．映像コンテンツを引用したブログエントリーの集合を本論文ではビデオブログと呼ぶ．想定される利用方法としては，ビデオコンテンツの紹介を目的とした記事の記述があげられる．

ビデオの任意のシーンの内容を表すサムネイル画像，そのシーンへのリンク及びそのシーンに対応するユーザコメントから成る段落をシーン引用パラグラフと呼び，ビデオブログエントリーは一つ以上のシーン引用パラグラフから構成される．シーン引用パラグラフの書式を統一することで，アノテーションの解析を行い易くする意図がある．

5.1 引用シーンの選択

ユーザはコンテンツを閲覧する際，自身にとって興味のあるシーンに対してシーンコメントアノテーションやシーンボタンアノテーションなどの何らかのアノテーションを施す．しかしながら，それらのアノテーションは会話的なコメントである，コメント情報が含まれていない等，必ずしもアノテーションとして優れているとはいえない．そこで，システムはこれらのアノテーションを施したシーンをビデオブログエントリーの執筆のための引用シーン候補としてユーザに提示し，ユーザにこれらの候補を元にしたビデオブログエントリーの執筆を促す．これにより，シーンアノテーションの投稿履歴から，段階的にユーザへより説明的な記述が期待できるブログ執筆を促し，より質の高いアノテーションの取得を目指す仕組みである．

5.2 ビデオブログエントリーの編集

ユーザが，ブログなどで通常のエントリーを書くのと同様に，一般的な Web ブラウザを用いてビデオブログエントリーの編集が可能になる仕組みを提案する．

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，連続する映像シーンを引用するのに適した編集インタフェース (図 5) である．これは，引用シーンをショット単位で時間的に展開させることで引用シーンの時間範囲を伴う修正・変更が可能であり，より正確にシーンを選択することが可能なインタフェースである．具体的には，シーン伸縮ボタンを押して引用シーンを時間的に前後に伸縮させることによって，正確に引用シーンを提示・選択可能であり，対応するコメントの編集も可能である．これは，シーンの流れやストーリを対象としたビデオブログエントリーを記述するのに適したインタフェースであると同時に，より詳細なアノテーションを施すためのツールでもある．連続する映像シーンとブログエントリー上の対応するパラグラフ上のコメントとを関連付けることを連続シーン引用アノテーションと呼ぶ．

二つ目は，複数の非連続な映像シーンを引用するのに適した編集インタフェース (図 6) である．過去にユーザが施したシーンコメントアノテーションやシーンボタン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）を対象としたビデオブログエントリーを記述するのに適したインタフェースである．また，映像シーン検索機能と併用することで，他のコンテンツのシーンの引用も可能である．これによって取得されるアノテーションを非連続シーン引用アノテーションと呼ぶ．

ユーザはこの二つのインタフェースを使い分けながらビデオブログエントリーを作成可能である．

図 5 連続シーン引用アノテーションインタフェース
Fig. 5 Continuous Scene Quotation Interface

図 6 非連続シーン引用アノテーションインタフェース
Fig. 6 Discrete Scene Quotation Interface

ビデオブログエントリーは HTML 文書として表現され，任意のブログサイトに投稿可能であると同時に，アノテーションデータベースに蓄積される．

6. アノテーションの解析

本システムでは，コメントアノテーションやシーン引用アノテーションを，なるべく情報劣化がない形式で蓄積する．そのため，本研究で意味するところのアノテーションはユーザコメントの列挙にすぎず，それ自身が機械によって理解可能な情報とは限らない．つまり，本研究によって取得されたアノテーションを用いたアプリケーションを構築するためには，アノテーションを解析し，機械が理解可能な情報に変換する必要がある．そこで，本章では 3 つの視点からアノテーションを解析する手法を提案する．一つは，アノテーションのテキスト情報からコンテンツの意味内容を表す情報の抽出を行う仕組みであり，具体的には，映像コンテンツ全体およびシーンの内容を表現するキーワード（一般にタグと呼ばれる）の抽出を目指す．二つ目は，アノテーションや映像シーンの各々の重要性の計算手法の提案であり，三つ目は，各々のアノテーション間やシーン間の関連性についての考察である．これらは，アノテーションに基づく応用を実現するために重要な情報である．

6.1 タグの抽出

アノテーションとして付与されたテキストからコンテンツやシーンの内容を表現するキーワードの抽出を行う．コンテンツと対応付けられたキーワードをタグと呼ぶ．特に，コンテンツ全体の内容を表現するタグをコンテンツタグといい，シーンの内容を表現するタグをシーンタグと呼ぶ．コンテンツタグ・シーンタグともに以下の手法によって抽出する．まず，それぞれの自由コメントを形態素解析器茶筌[17] を用いて形態素に分割する．それぞれの形態素から，名詞・動詞・形容詞・形容動詞・未知語を抽出する．ただし，代名詞や非自立名詞・非自立動詞は除外し，未知語は固有名詞として扱った．さらに一般的に不要語と判断可能な形態素（たとえば，する，ある，なる，できる，いる，等）も除外した．それぞれの形態素の基本形をタグとする．

6.2 アノテーションとシーンの重み

アノテーションやシーンの重みの計算手法を議論する．ここでいうアノテーションの重みとは，そのアノテーションが対象となる映像シーンの内容をどれだけ的確に，かつ，信頼性が高く表現しているかを示す指標であり，シーンの重みとは，そのシーンがその映像の中でどれだけ重要なシーンであるかを示す指標である．本来，重要なシーンとは状況や嗜好・目的に応じて変化する[16] ものである．しかしながら，PageRank[10] のように状況や目的を考慮しない重みづけによる検索システムであっても一定の成果をあげており，本論文では PageRank の概念，つまり，より参照されるシーンほど重要であるという指標に基づいて重要度を算出する．

具体的には，アノテーションの重みは，アノテーションの対象粒度，アノテータの信頼性，アノテーションタイプの信頼性から推定する．つまり，信頼できる人がより正確にアノテーションを作成できるツールを用いて，より粒度の細かい対象（コンテンツよりもシーン，長いシーンよりも短いシーン）に対するアノテーションを付与した場合に，より高い重みを与える．本来ならばアノテーションの意味内容を加味したアノテーションの重み付けをすることが望ましいが，本論文では意味内容を考慮したテキスト解析は一般に困難であるため見送っている．

また，映像シーンの重みは，より多くの，よりアノテーションの重みが大きいアノテーションから参照されているシーンほど重要であると仮定し，それぞれのシーンを参照するアノテーションの重みの合計がその映像シーンの重みであるとする．

具体的なアルゴリズムの提案と妥当性の検証は，十分なデータが不足している，コンテンツの種類やコミュニティに依存しやすいため検証が困難などの理由から，今後の課題とし，本節ではアノテーションとシーンの重みの計算手法のコンセプトのみを提示する．

6.3 アノテーション構造の活用

映像シーンに対するコメントアノテーションは，図 7 のように，対応する映像シーンとコメントとを「シーンコメントアノテーション」というラベルのついたグラフで表現される．コメントは映像シーンに関する情報を含んでいる場合が多く，映像シーンに対するアノテーションとして利用可能である．その一方，ビデオブログエントリーは，図 8 のように，引用した映像シーンとブログエントリーのパラグラフとを「シーン引用」というラベルのついたグラフで表現され，他のシーンやコンテンツ，ブログエントリーとの何らかの関連性の抽出が期待できる．

具体的には，連続シーン引用アノテーションによって選択された連続するショットからなる引用シーンでは，それに対応するコメント内容という観点に基づきシーンの連続性があるとみなすことができる．また，非連続シーン引用アノテーション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，シーンの関連性があると考えられる．さらに，一つのビデオブログエントリーで複数のコンテンツを同時に引用した場合，そのビデオブログエントリーの内容に基づいて，これらのコンテンツの意味的な関連性があると捉えることが可能になる．複数のコンテンツを引用したビデオブログエントリーの例としては，CG アニメーション「ノラネコピッピ 1 話」とその元になった実写映像である「ノラネコピッピのモデルになった猫♪」を同時に引用し比較する記事などである．

本システムにより，Web と映像コンテンツの垣根を越えた引用に基づく詳細なネットワークを形成する．これによりブログネットワークと映像コンテンツを統合することが可能になる．ブログと映像コンテンツの統合されたネットワークでは，コンテンツを扱う粒度がコンテンツ/エントリー単位から映像シーン/パラグラフ単位へとより詳細になり，コンテンツに関連するコミュニティが共有サイト内から Web 全体に拡大されている．さらに，コンテンツ間のリンクをナビゲーションのための 1 方向的な Hyperlink から引用に基づく意味的な双方向リンクへと拡張させることができる．これにより，我々の提案する仕組みはコンテンツに付随する様々な知識を抽出するためのフレームワークとして機能し，それによって収集されるデータは検索やコンテンツ推薦などの様々な応用のための基礎的データとして利用されることが期待できる．

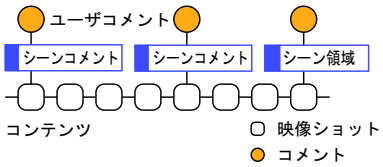

図 7 映像シーンへのアノテーションのモデル
Fig. 7 Video Scene Annotation Model

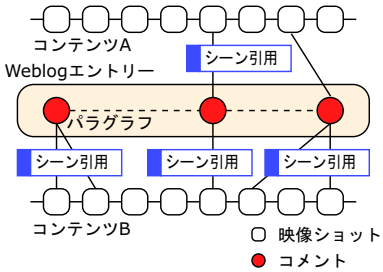

図 8 映像シーン引用に基づくアノテーションのモデル
Fig. 8 Annotation Model Based on Quotation of Video Scenes

7. 実験と評価

我々が提案したコミュニケーションを目的としたアノテーションから，検索等の応用に有用な情報がどれくらい取得可能であるかを検証するために，本論文で提案した Synvie☆の公開実験を行った．2006 年 7 月 1 日から公開を開始し，2006 年 10 月 22 日までの期間において収集されたデータに基づき評価を行う．この期間に，登録ユーザ数 97 人，投稿コンテンツ 94 個，1 コンテンツあたりの平均メディア時間は 321.5 秒，総閲覧数は 7318 回に達した．収集されたアノテーションは，表 1 で示すように計 4768 個であった．

コンテンツコメントアノテーションが YouTube 等の従来システムで実用化されているアノテーション，シーンコメントアノテーションが iVAS 等の従来システムによって取得されるアノテーションと捉え，本論文ではこれらに加えてシーン引用アノテーションを提案している．これらのアノテーションタイプの違いに

☆ http://video.nagao.nuie.nagoya-u.ac.jp/

表 1 公開実験によって取得されたアノテーション
Table 1 Result of Open Experiment

対象単位	行為	型	アノテーションタイプ	取得数
コンテンツ	投稿	文	コンテンツコメント	40
シーン	投稿	ボタン	シーンボタン	3412
		文	シーンコメント	795
			シーン領域コメント	187
	引用		連続シーン引用	283
			非連続シーン引用	51

表 2 アノテーションタイプ毎の有効タグ率と有効タグ精度
Table 2 Effective Tag Rate and Accuracy in Each Type

アノテーション	形態素数 (平均)	有効タグ数 (平均)	有効タグ精度
シーンコメント	7.19	1.51	58.8%
シーン領域コメント	7.77	2.17	60.9%
連続シーン引用	25.8	5.96	60.0%
非連続シーン引用	23.0	4.74	53.4%
コンテンツコメント	15.3	0.85	11.1%

よるアノテーションの質と量を比較することによって，シーン引用アノテーションの有用性を示す．

7.1 タグに基づく分析

タグの評価を行うために，あらかじめ全てのタグ候補に関して，そのタグが対応するコンテンツやシーンの内容を直接表現しているかどうかに基づき，筆者がタグの分類を手作業で行った結果を表 2 に示す．有効であると判断されたタグを有効タグと呼び，一つのアノテーションに含まれる有効なタグで重複のないタグの数を平均有効タグ数という．シーンコメントアノテーションとシーン引用アノテーションはどちらも映像シーンに対するアノテーションであるが，前者の平均有効タグ数は 1.51 であるのに対して後者は 5.96 と 3 倍以上多い．どちらも一つの映像シーンを話題としたコメントであるため，シーン引用アノテーションの方が，より詳細な話題について記述していることが推定される．シーンコメントアノテーションやコンテンツコメントアノテーションなどよりも，ブログ上で記述されるアノテーションの方がより多くのタグが含まれている傾向がある．次に，機械的に抽出されたシーンタグのうち，どれくらいのタグがそのシーンの内容を的確に表現しているかどうかを示す割合として有効タグ精度という割合で評価する．これは，一つのアノテーションに機械的に除去できない，ノイズとなるタグがどれだけ含まれていないかを表す．すべての形態素をタグとした場合の有効タグ精度は平均 20%前後であるが，前述したタグの絞り込み手法を用いると表 2 で示すように 60%前後まで向上する．この数値は決して高いとはいえない．しかしながら，有効タグは対応するシーンやコンテンツに直接関連しているかどうかという基準で選別したために，有効タグ率には映像から派生した話題に関連しているタグは反映されていない．無作為に記述されたアノテーションでない限りは，そのコンテンツを閲覧して記述したという観点から何らかの関連性はあり，派生的に関連したタグも含めれば有効タグ率はこれよりも大きくなる可能性が高い．アノテーションタイプ別の傾向として，シーンコメントアノテーションやシーン引用アノテーションなど，対象単位が映像シーンとなるものの有効タグ率が高い．これらは，映像シーンというより粒度の細かい対象について議論しているため，コメントの内容が映像シーンの内容に影響を受けやすいためであるからと考えられる．

タグの分類を手作業で行うのには多大のコストがかかることが懸念される．我々は二種類の方式でこの問題を解決することを考えている．一つは，構文解析や意図解析などといったより高度な言語処理技術を用いる手法である．人的コストがかからないという利点がある一方，Synvie で取得されるような自由コメントに対してこれらの問題を適用することは非常に困難である．二つ目は，増田ら[21] が提案した，ユーザらによって協調的にタグを選別する手法であり，費用対効果の観点から有用性が確認されている．

7.2 アノテーションの主観的分類

収集されたアノテーションを評価するために，それぞれのアノテーションのコメント内容に対して，以下のとおり，アノテーションの意味に基づく分類を行った．

A 主にシーンの内容を説明・解説するコメント．

B 主にシーンに対する直接的な感想や意見などからなるコメントで，シーンに関連するキーワードが含まれるもの．

C 主に，シーンの内容から派生した話題に関するコメント．

D 感嘆符のみ，形容詞のみなど単独では内容を理解できないもの．あるいは，撮影手法，映像の品質に対する感想など，シーンの内容とは関係のない話題からなるコメント．

さらに A，B，C のカテゴリに関して，コメントの文章としての正しさに基づき，

X コメントに主語・述語・目的語が存在するなど，十分に内容を表現している．

Y 十分に内容を表現しているとはいえない．

のサブカテゴリに分類した．なお，分類は二人の評価者によって同時に行い，異なる意見が出た場合には話し合いによる調整を行った．

$A-X$ のアノテーションの例としては，朝顔の展示に対して映像撮影者が自身のブログで「名古屋式盆養切込みづくりの朝顔です。 蔓を伸ばさずに盆栽仕立てにしていてとてもユニークです。100 年の歴史があるそうです。」と記述したコメントのように，シーンの内容を的確に表現しており言語解析などを行うことによって，より多くの知識の抽出が期待できる．$A-Y$ のアノテーションの例としては，Web アプリケーションのデモ映像で画像のアップロードを行っているシーンに対する「画像のアップロード」というコメントのように，シーンの内容を表現しているキーワードを含んでいるが，十分に内容を表現しきれていないものである．$B-X$ は「私にとっての朝顔は、こういう蔓を上へ上へと伸ばしていくタイプです。」のようにシーンに対しての感想や意見を述べているものであり，$B-Y$ は「どれだけお菓子使うんだよ！笑」のような表現であり，共にシーンの内容に関するキーワードの抽出が期待できる．$C-X$ の例としては映像中に表示される URL のキャプションに対して「リサイクルトナー専門店のようです。 著作権フリーの CG, 音楽を製作されているみたいです。」．$C-Y$ の例としては「長尾先生といえば，アノテーションの研究」などである．これらは，関連する話題について記述しており，必ずしも映像シーンの内容を直接的に表現していないが，シーンに関する補助的な情報としての利用が期待できる．D の例としては「すごっ！」や「キター」など，単独では意味を成さないコメントや，「なんでこの回だけ映像がぶれてるのでしょう？ウィンドウズメディアエンコーダーという無料ソフトで、ノンインターレス化できるので是非。」など映像の品質に関する話題などが含まれる．

アノテーションタイプごとにカテゴリ分けし，集計したものを図 9 に示す．

カテゴリ $A-X$ に該当する，コンテンツの内容を説明・解説するためのブログエントリーは，コンテンツ投稿者自身によって執筆される事例が多く含まれた．これは，自分の投稿したコンテンツを広くいろいろな人に見てもらいたいがためであると推察される．

7.3 考　　察

まず，アノテーションの量の観点から考察する．ここで，アノテーションの量は，そのアノテーションを付与する手軽さや扱いやすさに関係していると仮定する．表 1 で示すように，従来型のコンテンツ全体に対するコンテンツコメントアノテーションよりもシーンコメントアノテーションの方が投稿数が多いため，シーンコメントアノテーションはより手軽なアノテーション

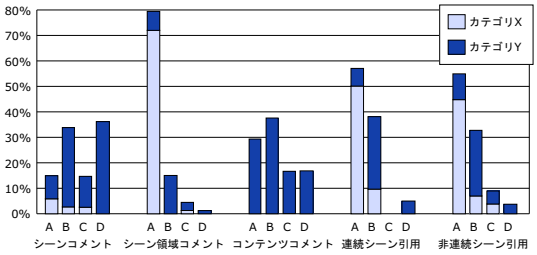

図 9 アノテーションタイプ毎のアノテーションの質の比較
Fig. 9 Quality of Annotation in Each Method

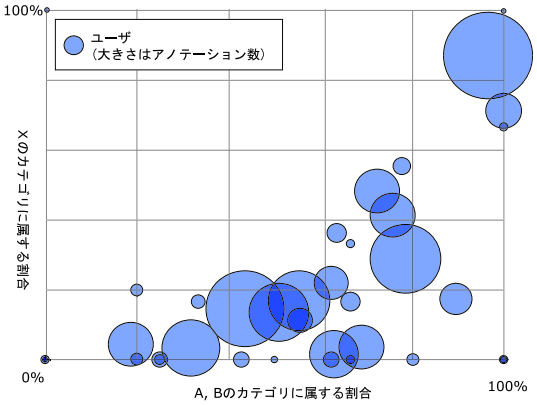

図 10 アノテーションを施した数及び品質に基づくユーザの分布．一つの円が一人のアノテータにあたり，円の大きさが投稿したアノテーションの数にあたる．右上に行くほど質の高いアノテーションを施したユーザである．
Fig. 10 Quality of Annotation in Each User

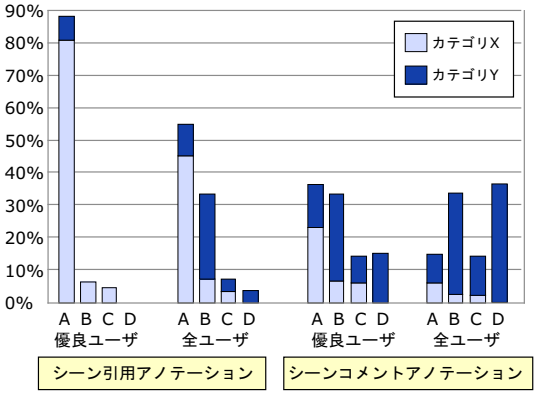

図 11 アノテーションタイプ及びユーザ毎のアノテーションの質の比較．優良ユーザとは，サブカテゴリ X に属するアノテーションを施した割合が多い人，上位 30%を示す．
Fig. 11 Quality of Annotation in Each Method and User

であったと推察できる．一見すると，コンテンツ全体に対するアノテーションの方が，シーンを選択する手間が無い分手軽であるように感じられるが，シーンに対するアノテーションの方が，注目対象を限定しているため，他の閲覧者と話題を共有し易く比較的短いコ

メントで内容を記述できる，些細な問題や話題でもコメントを投稿し易いなどの理由から，より手軽に投稿可能であるためだと推察できる。

次に，アノテーションの質の観点から考察する．厳密な質の定義は応用に依存するが，ここでは，コメント内容の品質が高くシーンの内容を的確に表現し，引用シーンに関連するキーワード等を含んでいるものとする．具体的には，$A > B > C > D$ の順で質が高く，また，サブカテゴリ X の方が Y よりも質が高いものとする．ただし，カテゴリ C に属するアノテーションは直接的にシーンに関係しているとはいえないコメントであっても，シーンから派生した情報であるため無関係とはいえない．むしろ，MPEG-7 などの通常のアノテーションからでは得ることが困難な重要な情報が隠れている可能性があり，決して無視することはできないと考えている．

これらの観点からみると，図 9 で示すように，コメントアノテーションに比べて，シーン引用アノテーションの質の方が高い．特に，シーンコメントアノテーションにおいてサブカテゴリ X に属する割合は 11%なのに対し，シーン引用アノテーションは 59%になり，より正確な文章が記述されていること，また，シーンコメントアノテーションにおいてカテゴリ D に属する割合が 36%も存在しているのに対して，シーン引用アノテーションの場合は 4.8%であるなど，無関係なコメントや “荒し” と呼ばれるコメントが少ないなどの点で，シーン引用に基づくアノテーションの方がより質が高い傾向があるといえる．

つまり，アノテーションの質や量はアノテーションタイプに依存する．これは，閲覧者が映像を見ているという前提が成り立ち，その場限りのコミュニケーションを目的としたシーンコメントアノテーションよりも，映像コンテンツを閲覧しているとは限らない不特定多数に向けたブログエントリーの執筆を目的としたシーン引用アノテーションの方がより丁寧な文章を記述する傾向があり，より質の高い情報を記述していると捉えることができる．また，掲示板よりもブログの方が一般的により良い文章が書かれている現状を反映した結果ともいえる．一見面倒で操作が多いアノテーションも，ブログを書く等といった人間の自然な日常活動の一部として取り込む事ができれば，十分な質と量を伴うアノテーションの取得が可能になることが分かる．

次に，アノテーションと人との関連性を考察する．図 10 に示すように，良いアノテーションを施す人もいれば，そうでない人もいる．つまり，人に応じてアノテーションの質や量は異なり，ばらつきがある．そこで，サブカテゴリ X のアノテーションを付与した数の割合が多い上位 30%のユーザ，つまり良いアノテーションを投稿する割合が多い人を優良ユーザと定義する．図 11 に示すように，優良ユーザがシーン引用アノテーション方式を用いて施したアノテーションのうち 80%が一番質の高いカテゴリである $A - X$ に分類される．これは，シーン引用アノテーション全体の平均の 45%や，質の高いアノテーションを施した上位 30%の人がシーンコメントアノテーションを用いて付与した平均 22%よりも圧倒的に多い．

つまり，アノテーションの量と質は人にもアノテーションタイプにも依存する．逆に言えば，人やアノテーションタイプが，アノテーションの質の推定パラメータの一つとして利用することが可能になる．具体的なアルゴリズムは現状ではデータが十分揃っていないため今後の課題としたいが，学習アルゴリズムを用いてパラメータを動的に決定することを検討している．

次にコンテンツの内容とアノテーションの量や質の関係について述べる．縦軸をアノテーション数，横軸をコンテンツとしてアノテーション数順に並べると図 12 のようになり，また，94 のコンテンツのうち，上位 12 個でおよそ半分のアノテーションの投稿数を占めている．この結果から面白いコンテンツほどより多くのアノテーションが投稿されることが分かる．次に，アノテーションの多いコンテンツとそれ以外のコンテンツでのアノテーションの質について議論する．アノテーションの多い上位 12 件のコンテンツを人気ありとし，それ以外を人気なしとした場合の，それぞれのアノテーションタイプ別の割合を図 13 に示す．ブログ引用の場合は人気ありのコンテンツの方がより質の高いアノテーションの傾向が認められるが，有意な差とまでは言えない．そのため，アノテーションの質はアノテーションの投稿数に依存しているとは認められない．

まとめると，YouTube 等で実用化されているコンテンツコメントアノテーションよりも，筆者らが iVAS で提案してきたシーンコメントアノテーションの方がより多くのアノテーションを収集することが可能であり，また，シーンコメントアノテーションよりも，本論文で提案したシーン引用アノテーションの方がより質の高いアノテーションの収集が期待できる．シーンコメント / シーン引用アノテーションを併用することによって，バリエーションに富んだ質・量とも高いアノテーションの収集が可能になる．これにより，多くのアノテーションが集まった場合は，評価が高いユーザが施したシーン引用アノテーションを重視し，あまり

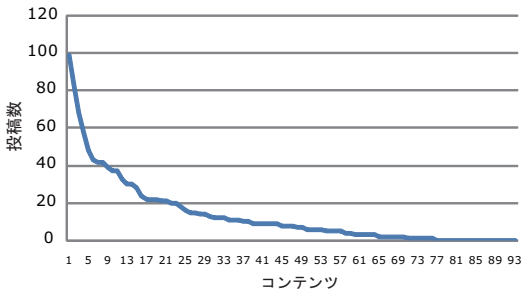

図 12 アノテーションの投稿数とコンテンツの関係
Fig. 12 Relation Between Annotations and Contents

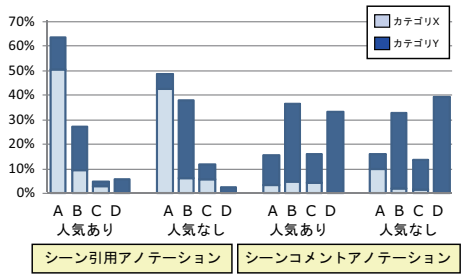

図 13 アノテーションの投稿数とアノテーションの質の関係
Fig. 13 Quality of Annotation in Each Content

集まらなかった場合は，シーンコメントアノテーションの情報も活用するなど，場合によって使い分ける事が可能になる．

8. アノテーションに基づく応用

本実験によって取得されたアノテーションに基づく応用の例として，ビデオシーン検索システムなどを提案する．具体的な応用を試作することによって，本実験によって収集されたアノテーションの有用性を示す．なお，対象コンテンツの数が１００個前後と少ないこと，アノテーションの量は時間とともに増えていき，それが応用の精度や質に直結すること，母体となるコミュニティやコンテンツに強く依存することなどから定量的な評価が困難であるため，詳細な評価は今後の課題とする．本論文ではアノテーションに基づく応用の可能性について言及することに留めておく．

8.1 映像シーン検索

映像シーン検索とは，映像をコンテンツ単位ではなくシーン単位で検索しようとする仕組みである．我々の手法の特徴は，アノテーションから抽出されたタグを検索することによって，ビデオシーンを検索しようとしている点である．

具体的な検索プロセスは以下のとおりである．ユーザは，目的のシーンを検索するために，一つないし複数の検索クエリをタグ形式で入力する．それらのタグと一致するタグを含むアノテーションの検索を行い，一致したアノテーションをコンテンツ毎に列挙する．一致したアノテーションに対応するシーンを検索結果候補とする．

一つのコンテンツ内に多くのアノテーションが一致した場合は，検索結果候補が膨大かつ時間軸上で細切れになる危険がある．そこで，対象となるシーンが連続する，あるいは時間的に近い場合は類似するシーンである可能性が高いと考え検索結果候補を統合する．逆に，一致したアノテーションの数が少なく，また，分散しており，検索結果のシーンを特定できない場合はコンテンツ全体を検索結果候補とする．検索結果候補内に属するアノテーションの重みの合計が，その検索結果候補の重みとする．このような仕組みにより，アノテーションが多数存在する場合にも，アノテーションが少量しか存在しない場合にも，ある程度対応可能になる．検索結果候補の重みに基づき，検索結果候補のランク付けを行う．

検索結果候補の内容を理解するために，シーンの内容を表現するサムネイル画像を提示することは有効である．サムネイル画像は，検索結果候補内のアノテーションに関連付けられているシーンに属するサムネイルを候補とする．ただし，サムネイル画像が一定個数以上存在する場合には，そのサムネイル画像が属する映像シーンの重みに基づいて絞る．

ビデオシーン検索システムのインタフェースを図 14 に示す．

検索が成功する例としては，検索したいシーンに的確なキーワードを含むアノテーションが存在する場合である．逆に，検索が失敗する例としては，検索したいシーンに的確なキーワードが含まれないなど，アノテーションの量が不足している場合が考えられる．しかしながら，人気のあるシーンやコンテンツには，より多くのアノテーションが集まりやすく，また，人気のあるシーンほど検索ニーズが高い，このようなシーンやコンテンツには自然にアノテーションが増えていくことが考えられる．すなわち，ある程度の時間が経過すれば，この問題は解決される可能性が高い．また，同じ内容を異なるタグで表現している場合にも検索に失敗する．その場合は，シソーラスを用いて類義語や語彙の上位概念・下位概念の関係を考慮する必要がある．

9. おわりに

本論文では，映像シーンへのアノテーション，映像シーン単位でのコンテンツの引用に基づくブログエン

図 14 ビデオシーン検索システム
Fig. 14 Video Scene Retrieval System

トリーからのアノテーション取得方法の提案，コミュニケーションに特化した具体的なインタフェースの提案と公開実験に基づく評価を行った．これにより，それぞれのアノテーションタイプによって得られるアノテーションの傾向をアノテーションの量と質の観点から分析を行い，それぞれのアノテーションに特有の傾向が見られることが分かった．特に，関連するブログエントリーから情報を抽出することが質の高いアノテーションを抽出する手助けになることが示せたことが有用であると考えている．これは，シーンコメントアノテーションが掲示板文化を引き継いでいるのに対して，シーン引用アノテーションはブログ文化を引き継いでいることを反映していると考えられる．また，これらのアノテーションは，二つの観点により映像の構造的・意味的情報も抽出可能である．一つは，コンテンツを引用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，引用によって複数のコンテンツ間の意味的な関係の抽出が期待できる．

今後の課題として，7 章 1 節の終わりに述べたタグ選別の自動化に関する問題や，アノテーションに基づく他のアプリケーションの開発が挙げられる．アプリケーションの例としては，ビデオ推薦システムやビデオスキミングシステムを想定している．我々が提案するビデオ推薦システムとは，映像と同期して関連性のある他のコンテンツのサムネイル画像とキーワード，及びその根拠となるビデオブログエントリーを表示し，関連するビデオの推薦を行うシステムである．本システムでは，複数コンテンツを同時引用したビデオブログエントリーの内容に基づく，統計情報に頼らない詳細なコンテンツ推薦を実現している．ビデオスキミング[15] とは，映像の重要なシーンのみを通常の速さで再生し，それ以外のシーンを早送りで再生する仕組みであり，映像の内容を短時間で把握するのに適している．具体的には，映像シーンの重要度を基にして，映像シーンの選別を行うことを考えている．

謝　　辞

本研究は独立行政法人情報処理推進機構 (IPA) による 2005 年度上期未踏ソフトウェア創造事業の支援を受けた．

参 考 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) Beged-Dov, G., Brickley, D., Dornfest, R., Davis, I., Dodds, L., Eisenzopf, J., Galbraith, D., Guha, R., MacLeod, K., Miller, E., Swartz, A. and van der Vlist, E.: *RDF Site Summary (RSS) 1.0*, RSS-DEV Working Group, http://web.resource.org/ rss/1.0/spec (2001).
3) Benjamin and Trott, M.: *mttrackback - TrackBack Technical Specification*, movabletype.org, http://www.movabletype.org/docs/ mttrackback.html (2002).
4) Davis, M.: An Iconic Visual Language for Video Annotation., *Proceedings of the IEEE Symposium on Visual Language*, pp. 196–202 (1993).
5) Dowman, M., Tablan, V., Cunningham, H. and Popov, B.: Web-Assisted Annotation, Semantic Indexing and Search of Television and Radio News, *Proceedings of the The 14th International World Wide Web Conference 2005 (WWW 2005)*, pp.225–234 (2005).
6) Hoffman, P. and Bray, T.: *Atom Publishing Format and Protocol (atompub)*, http://www.ietf.org/html.charters/atompub-charter. html (2005).
7) ISO: *Information Technology - Multimedia Content Description Interface(MPEG-7), ISO/IEC 15938:2001*, International Organization for Standardization(ISO) (2001).
8) Nagao, K., Ohira, S. and Yoneoka, M.: Annotation-Based Multimedia Summarization and Translation, *Proceedings of the Nineteenth International Conference on Computational Linguistics (COLING-02)*, pp.702–708 (2002).
9) 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).
10) Page, L., Brin, S., Motwani, R. and Winograd, T.: The PageRank Citation Ranking: Bringing Order to the Web (1998).
11) Parker, C. and Pfeiffer, S.: Video Blogging: Content to the Max, *IEEE Multimedia*, Vol.12, No.2, pp.4–8 (2005).
12) Smith, J.R. and Lugeon, B.: A Visual Annotation Tool for Multimedia Content Description, *Proceedings of the SPIE Photonics East, Internet Multimedia Management Systems*, pp. 49–59 (2000).
13) Wactlar, H.D., Kanade, T., Smith, M.A. and Stevens, S. M.: Intelligent Access to Digital Video: Informedia Project, *IEEE Computer*, Vol.29, No.5, pp.140–151 (1996).
14) 山田一穂，宮川　和，森本正志，児島治彦：映像の構造情報を活用した視聴者間コミュニケーション方法の提案，情報処理学会研究報告，Vol.2001-GN-43, pp.37–42 (2001).
15) 是津耕司，上原邦明，田中克己：映像の意味的構造の発見，情報処理学会論文誌，Vol.41, No.1, pp.12–23 (2000).
16) 梶　克彦，長尾　確：楽曲に対する多様な解釈を扱う音楽アノテーションシステム，情報処理学会論文誌， Vol.48, No.1, pp.258–273 (2007).
17) 奈良先端科学技術大学院大学自然言語処理学講座：形態素解析システム茶筌，http://chasen.aist-nara.ac.jp/ (2003).
18) 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).
19) 山本大介，増田智樹，大平茂輝，長尾　確：Synvie:映像シーン引用に基づくアノテーションシステムの構築とその評価，インタラクション 2007, pp.11–18 (2007).
20) 山本大介，清水敏之，大平茂輝，長尾　確：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，情報処理学会第 58 回グループウェアとネットワーク研究会，pp.13–18 (2006).
21) 増田智樹，山本大介，大平茂輝，長尾　確：オンラインアノテーションを利用したビデオシーン検索，第 21 回人工知能学会全国大会 講演論文集 (2007).

(平成 12 年 2 月 4 日受付)
(平成 12 年 5 月 11 日採録)

山本　大介（学生会員）
2003 年 名古屋大学大学院情報科学研究科メディア科学専攻修士課程修了, 2003 年-現在 名古屋大学大学院情報科学研究科メディア科学専攻博士課程. 日本学術振興会特別研究員. 2005 年 度上期 IPA 未踏ソフトウェア創造事業スーパークリエイター認定. 2006 年 情報処理学会 CS 領域奨励賞. Web と映像に関する研究に従事.

増田　智樹
2007 年 名古屋大学工学部電気電子・情報工学科卒業, 2007 年-現在 名古屋大学大学院情報科学研究科メディア科学専攻修士課程.

大平　茂輝（正会員）
2000 年 早稲田大学理工学研究科情報科学専攻修士課程修了, 2003 年 早稲田大学理工学研究科情報科学専攻博士課程単位取得退学, 2001-2003 年 早稲田大学理工学部情報学科助手, 2003-2006 年 名古屋大学情報メディア教育センター助手, 2006 年-現在 名古屋大学エコトピア科学研究所助手 (2007 年-助教).

長尾　確（正会員）
1987 年 東京工業大学総合理工学研究科システム科学専攻修士課程修了, 1987-1991 年 日本アイ・ビー・エム株式会社東京基礎研究所, 1991-1999 年 株式会社ソニーコンピュータサイエンス研究所, 1996-1997 年 米国イリノイ大学アーバナ・シャンペーン校客員研究員, 1999-2001 年 日本アイ・ビー・エム株式会社東京基礎研究所, 2001-2002 年 名古屋大学工学研究科助教授, 2002 年-現在 名古屋大学情報メディア教育センター教授.