

# 映像と論文の部分引用関係に基づく映像シーン検索の高度化

棚瀬 達央<sup>†</sup> 山本 圭介<sup>††</sup> 石戸谷 顕太郎<sup>††</sup> 大平 茂輝<sup>‡</sup> 長尾 確<sup>††</sup>

<sup>†</sup>名古屋大学 工学部電気電子・情報工学科

<sup>††</sup>名古屋大学 大学院情報科学研究科 <sup>‡</sup>名古屋大学 情報基盤センター

## 1 はじめに

近年、多種多様な学術的な内容を含む映像コンテンツが、インターネットを用いて手軽に閲覧できるようになり、映像シーン検索や映像要約のような応用に関する技術が求められている。それらを実現するためには映像コンテンツの内容を理解するヒントとなるアノテーションを獲得する必要がある。我々は、Web コミュニティ活動からアノテーションを獲得し、映像シーン検索に利用する仕組み [1] を開発したが、得られるアノテーションには映像とあまり関係のない情報も含まれることがあるため、映像シーン検索の精度が低下してしまう問題があった。そこで、本研究では、学術的な映像コンテンツには関連する論文が存在すると思われる、研究活動において、映像と論文を共に閲覧することがあることに着目し、論文部分の文章と映像の部分であるシーンを関連付けることで、映像アノテーションとして論文の文章を収集する仕組みを提案する。この関連付けで得られた情報を映像と論文の部分引用関係と呼ぶ。

また、収集した映像と論文の部分引用関係から、検索キーワードとなるタグを抽出し、そのタグには検索に有用なタグが多く含まれることを、従来手法で得られるタグと比較することで検証した。また、この部分引用関係から得られる映像アノテーションを利用して映像シーン検索を行うシステムを開発した。

## 2 映像と論文の部分引用関係に基づく映像アノテーション

我々が開発してきた複数のシステムを改良し、それらを組み合わせることで、映像の部分であるシーンと論文部分のテキストを関連付けて、論文のテキストを映像アノテーションとして獲得する仕組みを開発した。

### 2.1 映像の部分の引用

映像の部分であるシーンを効率的に作成する仕組みを提供し、手軽に映像シーンを引用できるシステムを開発した。まず、公開されている映像コンテンツをシステムに登録し、図 1 のようなシーン引用インタフェースで参照する。

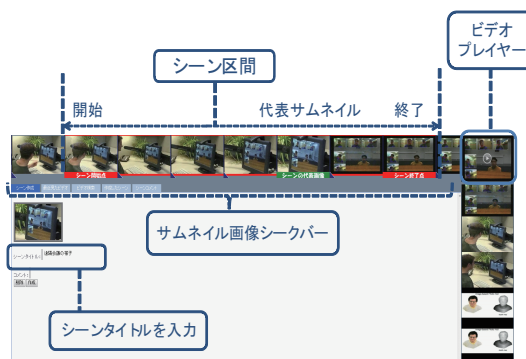


図 1: 映像シーン引用インタフェース

ユーザは対象となるビデオと、ビデオに同期して表示されるサムネイル画像を参照しながら、引用したいシーンの時間区間の決定を行う。サムネイル画像をマウスクリックによってシーンの開始、終了、代表フレームを決定する。最後に決定したシーンにタイトルを付けて、データベースに保存する。このインタフェースによってシーンというセグメント情報が蓄積され、それを引用することができる。

### 2.2 論文の部分の引用

論文の部分を作成するための仕組みを提供し、その論文部分を手軽に引用するシステム TDAnnotator を開発した。まず、公開されている論文を TDAnnotator に登録する。TDAnnotator では論文の任意の箇所を矩形により決定し、その箇所に対して翻訳やコメントなどのアノテーションを付与できる。図 2 に、アノテーションを作成するインタフェースを示す。ユーザは、自身で論文の箇所を決定して、アノテーションを付与し、それを利用して、論文部分を引用するだけでなく、すでに他のユーザによって作成されたアノテーションを検索して、それが付与された論文部分を引用することもできる。

### 2.3 知識活動支援システムを利用した関連付け

映像と論文を引用する仕組みとして、我々の研究室で開発された DRIP システム [2] を利用する。これは、個人の研究活動における、ユーザが行った文献調査やシステムの構築といった様々なタスクを記録し、それらの間のリンク情報と共に可視化する Web サーバ・クライアント型アプリケーションである。このシステムを、論文と映像シーンも扱えるように拡張し、映像シーンと論文の部分をリンク付けることで、部分引用関係の作成を可能にした。この関係に基づいて、映像シーンと関連付いた論文部分の文章を映像アノテーションとして獲得する。

Advanced Video Scene Retrieval Based on Co-Quotations of Video Scenes and Technical Documents

<sup>†</sup> TANASE, Tatsuo (tanase@nagao.nuie.nagoya-u.ac.jp)

<sup>††</sup> YAMAMOTO, Keisuke (kyamamoto@nagao.nuie.nagoya-u.ac.jp)

<sup>††</sup> ISHITOYA, Kentaro (ishitoya@nagao.nuie.nagoya-u.ac.jp)

<sup>‡</sup> OHIRA, Shigeki (ohira@nagao.nuie.nagoya-u.ac.jp)

<sup>††</sup> NAGAO, Katashi (nagao@nuie.nagoya-u.ac.jp)

Dept. of Information Engineering, School of Engineering, Nagoya University (<sup>†</sup>)

Graduate School of Information Science, Nagoya University (<sup>††</sup>)

Information Technology Center, Nagoya University (<sup>‡</sup>)

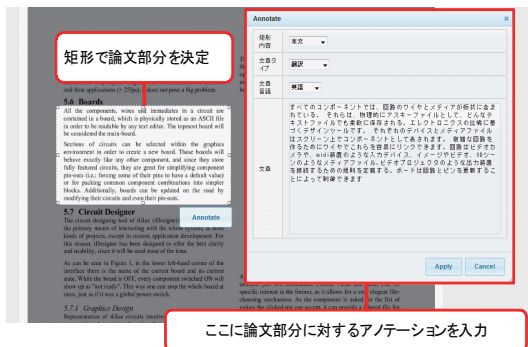


図 2: 論文部分へのアノテーションインタフェース

## 2.4 実験と考察

本研究では、アノテーションのテキストからシーンと関連するキーワード(シーntag)を抽出し、それを用いてシーン検索を行う。そこで、提案手法により作成した映像アノテーションである論文のテキストから抽出されるシーntagの量と質について検証するために、従来手法である映像シーンに対するコメントのテキストから抽出されるタグとの比較実験を行った。

このシステムを用いて映像シーンと論文のテキストのデータセットを作成し、作成した映像シーンに対するコメントを被験者実験により収集した。実験で使用した映像シーンには、被験者により、あらかじめ自然なコメントができそうなシーンとして選別された16シーンを用いた。計10人の被験者により合計92個のコメントが作成された。

1つのシーンに投稿された複数人によるコメントは1つのアノテーションとして扱い、含まれる名詞・複合名詞の異なり数(タグ数)のうちの特徴語として判定された単語の数を調べた。表に実験結果を示す。

表 1: 得られたタグ数と、それに占める特徴語の割合

	特徴語数	タグ数	特徴語の割合 (%)
論文部分	158	572	27.6
コメント	107	418	25.6

論文にはコメントよりも多くの特徴語が含まれていることが示された。また、論文の特徴語の質についてみると、論文には、コメントなどの口語の文章には含まれにくい専門用語が多く含まれていた。論文の文章からは従来手法であるコメントよりも、検索に有用な特徴的なタグが多く抽出されることになるので、これにより映像シーン検索の精度が向上すると考えられる。

## 3 映像シーン検索

提案手法により作成した映像アノテーションである論文のテキストから抽出されるタグに基づいて、映像シーンを効率的に検索するシステムを開発した。検索のトップページには、シーntagの一覧が表示され、その中から検索クエリを選択し、検索を行う。その結果、クエリのタグが関連付けられているシーンに対し

てスコアリングがなされ、検索結果としてシーンがランキング表示される。また、シーンごとのスコアの総計によるコンテンツのランキング表示も行える。コンテンツのランキング表示による検索結果(図3)では、シーン情報をブラウザ上で表示するための時間軸シークバーとサムネイル画像、シーntag、関連付けられている論文が一覧で表示される。

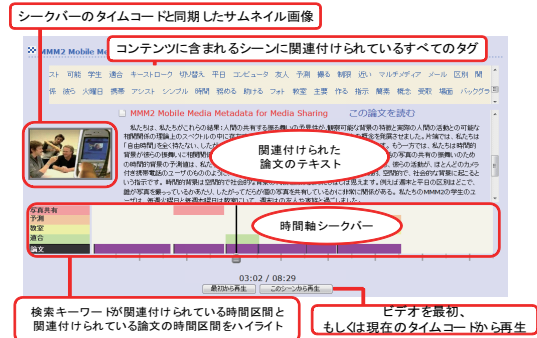


図 3: シーンに基づいて映像コンテンツを俯瞰するインタフェース

時間軸シークバーには、検索クエリのシーntagが関連付けられている時間区間と、そのコンテンツにおいて論文が関連付けられている全ての時間区間がハイライト表示される。シークバーを動かすことで、そのタイムコードに同期してサムネイル画像が切り替わることに加え、そのタイムコードを含むシーンに関連付けられている論文の文章を閲覧することによって、映像の内容を理解することができる。また、一覧中のタグをクリックすることでそのタグをシークバーに反映させることができる。さらに、関連付けられている論文の閲覧や、その論文が関連付けられている別の映像コンテンツを検索することも可能である。このインタフェースで、ユーザは、シーンに限定されない、任意のタイムコードからの映像視聴も可能である。

## 4 まとめと今後の課題

本研究では、研究活動における映像シーンの引用と論文部分の引用、さらにそれらの関連付けを行うシステムを開発し、その部分的な引用関係から映像アノテーションとしての論文の文章を獲得する仕組みを実現した。また、従来手法との比較実験によって、提案手法の有効性を確認した。今後の課題として、高度な言語解析によって論文のテキストとシーンとの間の意味的関係の推測によるタグの選別や、クラスタリングによるタグ間の関連性を、検索に反映させる仕組みの実現などが挙げられる。

## 参考文献

[1] Tomoki Masuda, Daisuke Yamamoto, Shigeki Ohira, and Katashi Nagao. Video Scene Retrieval Using Online Video Annotation. Lecture Notes on Artificial Intelligence (LNAI 4914: JSAI 2007 (K. Satoh, et al. Ed.) Springer-Verlag, pp.54-62, 2008.

- [2] 土田 貴裕, 大平 茂輝, 長尾 確, ゼミコンテンツの  
再利用に基づく研究活動支援情報処理学会論文誌,  
Vol.51, No.6, pp.1357-1370, 2010.