

コメントアノテーションに基づく講義コンテンツのシーン検索

笠嶋 公一朗† 山本 圭介†† 大平 茂輝‡ 長尾 確††
 †名古屋大学 工学部電気電子・情報工学科 ††名古屋大学 大学院情報科学研究科
 ‡名古屋大学 情報基盤センター

1 はじめに

近年、情報通信技術を用いた学習である e ラーニングに注目が集まっているが、その中で最も注目されているのが WBT(Web Based Training)[1]であり、その例としてオープンコースウェアが挙げられる。これは高等教育機関が提供する講義コンテンツをネットワーク上に無償公開し、学習支援に役立てる取り組みであるが、ブロードバンド化の恩恵により、講義映像の公開も可能になった。しかし、学習者がただ講義映像を見るだけでは、効率良く理解することは難しい。

そこで本研究では、講義スライドを介して行われる学習者間のコミュニケーションをシーンへのコメントアノテーションとして利用することで、コンテンツのシーン単位での検索を可能にする。先行研究として、映像のシーンに対するアノテーション手法とそれを利用したシーン検索に関する研究 [2] があるが、これらは一般的な映像に対する手法である。本研究は講義映像に限定し、講義専用のシーン検索を提案する。そしてシーン検索の結果を手がかりとしつつコンテンツ全体を俯瞰できるようにすることで、学習者の効率的な学習を支援することが可能になると思われる。

一般に、シーンは映像に対する時間情報であるため、シーン検索を行うには、講義映像に対しシーンを定義しなければならない。講義内容を把握している講師がシーンを決めることが理想であるが、それは保証されないため、システムがシーンを決める必要がある。そこで、講義スライドの表示時間情報を利用する。講義スライドには必ずしも時間情報は含まれていないが、講師によるスライドの切り替えタイミングを取得し、それを講義映像に関連付けることで時間情報を与えることができる。それにより、スライドの表示時間情報を講義映像のシーンの単位とみなすことができる。

2 講義コンテンツ

本研究における講義コンテンツは、講義映像と講義スライド、そして、講義スライドを介して行われる学習者間のコミュニケーションによって得られるコメントアノテーションを含んでいる。

2.1 コメントアノテーション

学習者が講義映像を閲覧しているときに理解が困難な内容が現れた場合、講義スライドの関連する部分を指定し、その部分に対して質問などのコメントを書き込めるようにする。そして、そのコメントを見た他の学習者が、そのコメントを選んで回答コメントを書き込むことができる。あるいは講義スライド内で表現の分かりにくい部分や回答コメントなどに対し、補足的なコメントをすることができる。このような学習者間のコミュニケーションによって得られるコメントのこ

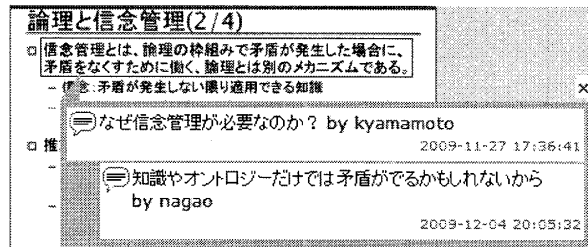


図 1: コメントアノテーション

とをコメントアノテーションという。図 1 はコメントアノテーションとそれが指しているスライドの部分を表したものである。

2.2 講義コンテンツにおけるシーン

講義スライドのアニメーション単位での表示時間情報を取得し、それをシーンの単位として利用する。それにより、表示切り替え時に新たに表示されたスライド内のテキストとその部分を指定して付けられたコメントアノテーションが、そのシーンに対して関連付けられる。

また、講義スライドのページ単位での表示時間情報も、より大きなシーンの単位として用いる。この場合、ページ全体を指定して付けられたコメントアノテーションが関連付けられる。

講義スライドは講義映像と同期的に表示可能なため、一般に映像の時間情報であるシーンを、講義スライドの表示時間情報で定義することができる。

3 講義コンテンツのシーン検索

シーン検索は、前述したシーンに関連付けられたテキスト情報を形態素解析し、生成されるシーンタグと呼ばれるタグを用いて行う。シーンタグは、生成元のテキストが関連するシーンに対して関連付けられる。

3.1 検索クエリが関連付けられるシーンの拡張

検索クエリ (ユーザが検索に使用したシーンタグ。複数の場合もある) に関連付けられるシーンを以下の順で拡張することで、クエリと直接には関連付けられていないが、何らかの関わりがあると思われるシーンを推測する。

- 検索クエリが直接関連付けられているシーン
- スライドテキストおよびコメントの共起情報に基づきシーンタグをクラスタリングし [3]、検索クエリ中のシーンタグが属するクラスタに含まれる他のシーンタグ (関連シーンタグと呼ぶ) と直接関連付けられているシーン
- (a)(b) によって得られたシーン (リンク元シーン) と以下の方法でリンクが結ばれているシーン (リンク先シーン)

- TFIDF に基づき各シーンの特徴語を抽出した結果、特徴語が一致するもの
- 1 つのコメントが複数のシーンを指す場合に、同時に指されたもの (図 2)
- 同一学習者によるコメントに共通のタグを含む場合に、それらのコメントの付与されたもの
- スライド内の画像が類似するもの

Video Scene Retrieval in Lecture Contents Based on Comment Annotations

† KASASHIMA, Koichiro(kasashima@nagao.nuie.nagoya-u.ac.jp)

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

‡ OHIRA, Shigeki(ohira@nagoya-u.ac.jp)

†† NAGAO, Katashi(nagao@nuie.nagoya-u.ac.jp)

Dept. of Information Engineering, School of Engineering,

Nagoya University (†)

Graduate School of Information Science, Nagoya University

(††)

Information Technology Center, Nagoya University (‡)

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

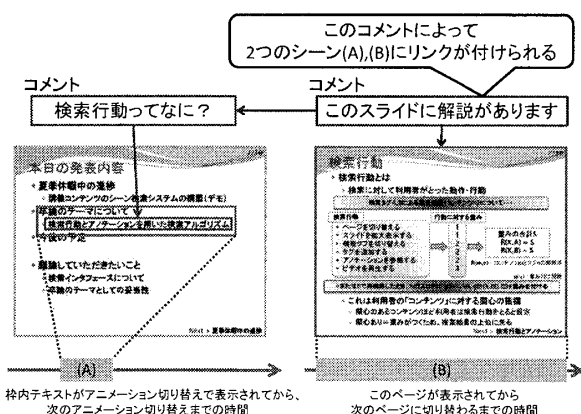


図 2: コメントの関連箇所が複数ある場合のリンク付け

表 1: シーンに対するスコアの算出方法

シーン	スコアの算出方法
(a)	シーンに含まれる検索クエリの数
(b)	関連シーンタグの数/クラスタの要素数
(c)	リンク元のシーンのスコア×リンクの強さ (リンクの付け方により決まる正定数。同シーン間に複数のリンクが付く場合、上限を 1 として加算される)

そして、関連付けられたシーンに対してスコアリングをする。スコアの算出方法を表 1 にまとめる。コンテンツのスコアは、コンテンツに含まれるシーンのスコアの総計によって得られ、スコアの高い順にコンテンツがランキング表示される。

3.2 シーン検索のユーザインタフェース

検索のトップページにはシーンタグの一覧が表示され、その中から検索クエリを選択し検索を行う (複数選択可能で、その場合 or 検索となる)。その結果、前述の方法で検索クエリが関連付けられたシーンのスコアリングおよびコンテンツのスコアリングが行われ、コンテンツがランキングされる。それが検索結果として返され、それぞれのコンテンツに対して、シーン情報をブラウザ上で閲覧するための時間軸シークバーと講義映像、講義スライドのサムネイル一覧、シーンタグとコメントアノテータの一覧が表示される (図 3)。

時間軸シークバーの上部は、検索クエリが関連付けられたシーンがハイライト表示される。(a) の場合はその検索クエリに割り当てられた固有の色が、(b) の場合は紫色が、(c) の場合は灰色が割り当てられており、その色が濃いほど、シーンのスコアが高いことを意味する。また、各シーンのスコアをスライドのページ単位で集計し、その値が高いほど、そのページのサムネイル枠の色がより赤くなり、ページ単位での検索クエリの関連の強さが視覚的に分かることができる。シークバーの下部は検索クエリに関係なく、各シーンに対するコメントの数をアノテータで区分けしてハイライト表示させた。「!」マークは最も直近についたコメントのあるシーンを表している。

時間軸シークバーの上部にあるハイライト部分をクリックすると、そのシーンに関連付いたスライドテキストおよびコメントの中から、対応するタグを含むものがポップアップ表示される。下部の場合、対応するアノテータがそのシーン (対応するスライドにあるいはその一部) に対して付与したコメントを、投稿日時の新しい順に並べてポップアップ表示させる (図 4)。

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映像とスライドサムネイル画像が切り替わる。映像プレイヤー部分をクリックすると映像の再生を行

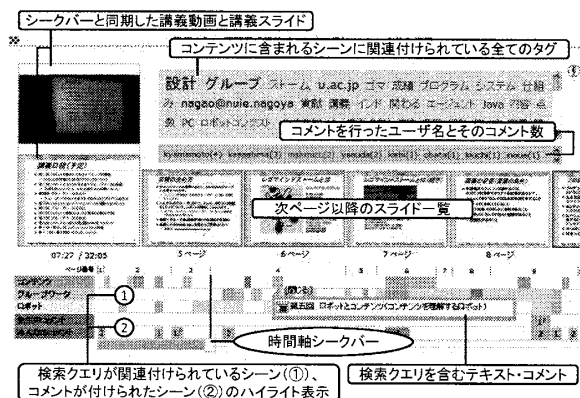


図 3: 検索のユーザインタフェース

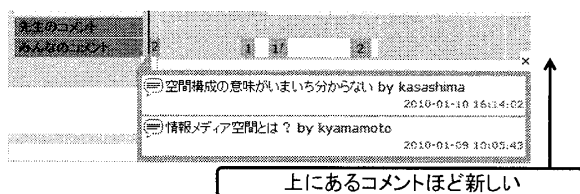


図 4: コメントのポップアップ表示

うことができる。これらの仕組みによって、シーン検索の結果のハイライト部分を手がかりとして、学習者がシーンを選択、閲覧でき、またコンテンツ全体を俯瞰しつつ、そのシーンがコンテンツ全体においてどういった位置付けなのか、どういった文脈の中で現れたシーンなのかを把握できる。それが新たな手がかりとなって、他のシーンの閲覧や文脈把握などを続けていくことにより、講義の理解を深めることが可能になるとと思われる。

4 まとめと今後の課題

本研究では、スライドを介した学習者間のコミュニケーションにより得られるコメントアノテーションを利用することで、コンテンツ全体を俯瞰しつつ、効率的にシーンを検索する仕組みを提案した。

今回定義したシーンの単位は講義スライドの表示時間情報を用いたものであったが、講義内容の意味的な単位を 1 つのシーンとして割り当ててくれる方が、講義コンテンツのシーンとしてより自然である。したがって、高度な自然言語解析や講義中の講師の動作や音声の解析などを行うことにより、1 つの講義を意味的な単位で分割する仕組みが必要となる。

それにより、検索クエリが関連付けられるシーンの拡張方法に再考の余地が生まれる。今回は主として自然言語処理に関する手法を用いたが、意味的な単位でシーン分割ができれば、各シーンにはその意味が属性として付加できるため、意味情報を用いてより高度にシーンの関連性を考慮することが可能になる。

参考文献

- [1] 中山幹夫: 教育課程における情報教育の課題と WBT(Web Based Training) の役割, 情報文化学会第 10 回全国大会論文集, pp.85-88 (2002).
- [2] Masuda, T., Yamamoto, D., Ohira, S. and Nagao, K.: Video Scene Retrieval Using Online Video Annotation. In: LNCS (LNAI), Vol. 4914, pp. 54-62. Springer (2008).
- [3] 松尾 豊, 石塚 満: 語の共起の統計情報に基づく文書からのキーワード抽出アルゴリズム, 人工知能学会論文誌, Vol.17, No.3, pp.213-227 (2002).