情報処理学会論文誌　データベース　Vol. 2　No. 2　71–85 (June 2009)

プレゼンテーションコンテンツのための シーンの意味的関係抽出とその応用

北山大輔[†1]　大谷亜希子[†2]　角谷和俊[†3]

近年，プレゼンテーションを録画したビデオと発表スライドを同期させたプレゼンテーションコンテンツ視聴環境が整ってきている．このような環境では，ユーザはコンテンツを蓄積したアーカイブを利用し，視聴したいコンテンツを取得することができる．大量のコンテンツから関連するシーンを抽出する方法としては，コンテンツの各シーンに含まれるキーワードを全文検索して抽出する手法が考えられる．その際，コンテンツ全体の視聴ではなく，コンテンツの一部分の視聴や，複数のコンテンツの部分を断片的に視聴する場合がある．しかし，単純にコンテンツから検索キーワードを含むシーンを抽出したのでは，文脈が分からないという問題がある．そのため，コンテンツを構成する区間の意味的関係を考慮し，提示する手法が必要である．本研究では，シーン間の意味的関係に基づく視聴や要約を実現するために，階層構造情報を持つスライドと，ビデオから抽出した発表者の発話内容のタイムスタンプが付与されたテキスト情報に基づきシーン間の関係を抽出する手法を提案する．さらに，提案した手法をもとにプロトタイプシステムを構築し，評価実験および考察を行う．

An Extracting Method of Semantic Relations between Scenes for Presentation Contents

Daisuke Kitayama,[†1] Akiko Otani[†2] and Kazutoshi Sumiya[†3]

It has become easier for users to view presentation content consisting of synchronized heterogeneous media such as videos and slides. Users simply retrieve the stored content they wish to view. In most cases, users are interested in viewing a specific portion of the content. Scenes extracted from a presentation, however, often do not help the user understand the content, even when those scenes contain the query keyword. A method is needed for extracting viewing intervals from content which are meaningful to the user. In this paper, we propose a method to extract scenes according to their relation. These relations are decided at the point of change between scenes. In this paper, we describe our method and subject it to evaluation.

1. はじめに

近年，スライドなどの資料と発表者を撮影したビデオを同期させたプレゼンテーションコンテンツ（以降，単にコンテンツと記す）を視聴するシステムが，実際の学会発表を蓄積することに運用されるなど広く活用されるようになってきている[1)]．蓄積されたコンテンツは，聴講できなかったユーザが後日に内容を知るために視聴する場合に有効であると考えられる．このとき，視聴者はコンテンツをより適切に理解したいという要求を持ちながら視聴することが考えられる．たとえば，あるプレゼンテーションを視聴する場合，プレゼンテーションのストーリを適切に把握することが重要である．このとき，コンテンツ中のあるシーンに対して，「より詳細な内容を述べている」というようなアノテーションが付与されると，理解の促進につながると考えられる．また，視聴しているシーン中で興味があるキーワードを指定することで，「そのキーワードに関する一連のシーン群」が抽出されると，より効率的に視聴することが可能であると考えられる．

本研究では，コンテンツはスライドとビデオの 2 種類のメディアによって構成されると仮定している．また，スライドの切替えタイミングによりシーン抽出を行うものとする．すなわち，スライドの 1 枚に対して 1 つのシーンが対応することになる．ここで，「シーン A よりもシーン B はキーワード k に関してより詳しい内容」のような，シーン間の相対的な意味的関係を定義する．このような，シーン間の意味的な関係を抽出することで，「指定シーンの内容に関する要約」や「シーン間の関係を明示した状態での効率的な視聴」といった各種のアプリケーションが構築可能であると考えられる．本研究では，シーン間の意味的な関係を，スライドやビデオという異なるメディアのメタデータのシーン間における変化により分析する方式を提案する．

以下，2 章では，本研究の概要と関連研究について述べる．3 章では，コンテンツの特性とメタデータについて説明する．4 章でシーン間の関係の判定手法について述べる．5 章で応用アプリケーションについて説明し，6 章でプロトタイプシステムを用いた本手法の評価

†1 兵庫県立大学大学院環境人間学研究科・日本学術振興会特別研究員 DC
Graduate School of Human Science and Environment, University of Hyogo
†2 NTT データネッツ株式会社第一開発統括部
The First development group, NTT DATA Net's CORPORATION
†3 兵庫県立大学環境人間学部
School of Human Science and Environment, University of Hyogo

© 2009 Information Processing Society of Japan

とその考察を行う．

2. 本研究の概要と関連研究

2.1 本研究の概要

ユーザがコンテンツを視聴する際，指定したキーワードが含まれるシーンのみではなく，興味を持ったシーンを理解するために必要なシーン群によって構成される区間を視聴したいという要求があると考えられる．また，抽出されたシーン群を効果的に理解するために，それらのシーン間の関係性を分析したいという要求があると考えられる．このため，指定されたシーンと関係するシーンを提示し，その関係を明示することが効果的であると考えられる．しかし，一般にコンテンツの構成は作成者によって異なるため，適切に関係するシーンが，時間的に前後して出現するシーンや指定したキーワードが含まれるシーンであるとは限らない．そのため，コンテンツの構造的な情報を用いて関係するシーンを判定し，シーン間に意味的な関係が存在するシーン群を抽出する必要がある．

コンテンツ中のシーンは，他のシーンと意味的に関係を持つ場合がある．たとえば，あるシーンで説明した内容を，後のシーンで詳しく説明している場合，これらのシーン間には**詳細化**という関係があると考えた．このように，あるシーンと意味的に関係するシーンを判定し，結合することでコンテンツから適切に視聴区間を抽出し，ユーザに提示できると考えた．シーン間の意味的な関係として，**詳細化**，**汎化**，**具体化**，**付加**の 4 種類を定義する．これらの関係は，2 つのシーンが持つ異種メディアのメタデータを用いて判定する．

図 1 に本手法の概念図を示す．本手法では，コンテンツに対して，以下の手順により意味的関係の抽出を行う．まず，任意に決定したシーンに対して，総当たりで他のシーンとの間における関係を決定する．具体的には，同一キーワードのスライドでの階層構造の変化，ビデオでの発言量の変化を用いて 2 シーン間の関係の判定を行う．このようにして判定した関係をアノテーションとしてコンテンツに付与し，プレゼンテーションアーカイブとして蓄積する．

アーカイブの利用時においては，アプリケーションごとに定義する型判定を行い，シーン間の意味的関係を抽出することで各種の機能を実現する．型判定とは，シーンが持つ異種メディアのメタデータは用いずに，意味的関係のみを用いる判定手法であり，2 シーン間の関係の組合せによって，複数シーン間の関係の判定を行う．ここで構築することが可能であるアプリケーションとしては，ユーザが指定したキーワードに関して特定の関係を持つシーンを検索するアプリケーションや，コンテンツを視聴する際にコンテンツ全体におけるシーン間の関係が把握できるように関係を可視化するアプリケーションなどが考えられる．本稿では，過去に行われたプレゼンテーションを効率的に視聴するアプリケーションとして，プレゼンテーションのスライド一覧画面から興味を持ったキーワードを選択するだけで，意味的な関係のあるシーン群を抽出し，さらにシーン間の関係を明示した視聴を行うプロトタイプシステムを構築し，意味的関係抽出の評価を行う．

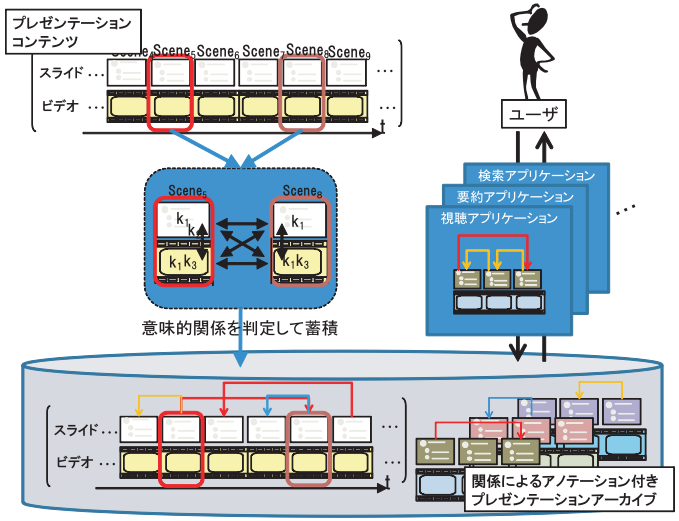


図 1　関係抽出とその応用の概念図
Fig. 1　Concept of extracting relations.

2.2 関 連 研 究

異種メディア統合コンテンツの特性の利用に関する研究として以下の研究があげられる．He ら[2] は，スライドを利用したプレゼンテーションコンテンツに対して，声の高さの解析，シーン提示時間，利用者の各時刻に対するアクセス頻度から重要箇所を抽出し，ダイジェストを作成する手法を提案している．吹野ら[3] は，実況コメントから映像のメタデータを生成し，高度な要約文であるニュース記事と照合を行うことで，ビデオ区間を抽出する手法を提案している．Smith ら[4] は，CC（Closed Caption），音声，画像ストリームの協調処理によりシーンの特徴を検出することで，映像から区間抽出を行う手法を提案している．小林らの提案する UPRISE[5] では，メタデータを用いて資料や映像の疎結合を行い，コンテン

© 2009 Information Processing Society of Japan

ツを蓄積し，提供する枠組みを提案している．村木ら[6]は，プレゼンテーション資料やビデオといった多様なコンテンツを，スライドと映像の同期情報やキーワードに対する適合度から構成されるメタデータを用いて統合する方式を提案している．Nakano ら[7]は，音声情報やレーザポインタ照射情報を用いた手法を提案している．従来の異種メディアコンテンツの特性を用いた研究では，コンテンツ全体から重要なシーンを判定し，要約を生成する手法が主として提案されている．本手法では，シーンの重要さを用いずにシーン間の相対的な関係を抽出することで要約に加えて，抽出されたシーンの関係を提示することが可能である．

映像の区間抽出に関する研究として Pradhan らの研究があげられる．Pradhan ら[8]は，断片的なビデオ区間集合を組み合わせ，新しい区間を生成するグルー結合を提案している．コンテンツ中のシーン間における部分的な情報から関係するシーンを抽出するという点で本研究と類似しているが，この手法は，映像という単一メディアを対象としたものであり，複数メディアからなるコンテンツのシーン間の関係を抽出する本手法とは対象が異なる．

戈ら[9]は，質疑応答時の資料操作を用いた，アーカイブコンテンツの再構成手法を提案している．北山ら[10]は，コンテンツのシーン間の関係を用いた区間検索方式について提案している．これらの研究における入力はシーン指定であり，シーン間の関係しか考慮することができていない．本方式ではシーン間の関係として，キーワード単位での関係を抽出することができ，より詳細に関係を判定することが可能である．また，シーンの関係を用いて，複数シーン間の関係を判定する型判定の概念を導入し，アプリケーションへ応用することを目的としている．

Web ページもハイパーリンクなどによって，意味的な関係を持つコンテンツであり，また HTML 構造のように，コンテンツの構造の特徴を用いて意味的な関係を抽出することが可能であると考えられる．そのため，Web ページを対象とした関連研究について述べる．本手法では，入力が行われたシーンにおける指定キーワードのスライドとビデオでの出現傾向によって，提示コンテンツを変化させる．小山ら[11]は，Web ページ内でキーワードの出現位置を考慮した検索方式を提案している．具体的には，同じキーワードでもタイトルに現れる場合と本文中に現れる場合とで，キーワードの意味を区別している．小山らは単一の Web ページの構造を対象として，キーワードの意味的な関係を抽出しているのに対し，本手法は 2 つのシーンの構造の違いを対象として，キーワードによるシーン間の意味的関係を抽出するため対象が異なる．

また本手法は，シーンごとに分割されたスライド集合，ビデオ集合を対象として意味的関係の抽出を行う．Yumoto ら[12]は，検索結果として質問キーワードに関して話題の広さと深さの両方を満たすページ集合を生成し，ユーザに提示する手法を提案している．Yumoto らはあるトピックを理解するために Web ページ集合を網羅という関係で抽出することを目的としているのに対し，本手法ではコンテンツに関し，あるシーンを理解するための基本となる意味的関係を定義し，関係を組み合わせることで各種のアプリケーションへ利用することを目的としている点で異なる．

3. コンテンツの特性とメタデータ

3.1 コンテンツのメタデータ

コンテンツは，スライドとビデオという同期した 2 種類のメディアから構成される．本章では異種メディアのメタデータについて説明する．スライドからはスライドに書かれた文字列とインデントによる構造化情報を得ることができ，ビデオからは音声のテキストが得られるものとする．スライドの切り替わりによって区切られる単位をシーンと呼ぶ．

本稿では，スライドのインデントによる階層構造を以下のように定めた．

- タイトルを階層の最上位である 1 とする．
- 本文の最上位を 2 としてインデントが深くなるたびに 1 階層ずつ増やす．
- 図や表などのインデント外となる要素はスライドの平均階層にあるとして扱う．

図 2 に本稿で扱うコンテンツの構造を示す．

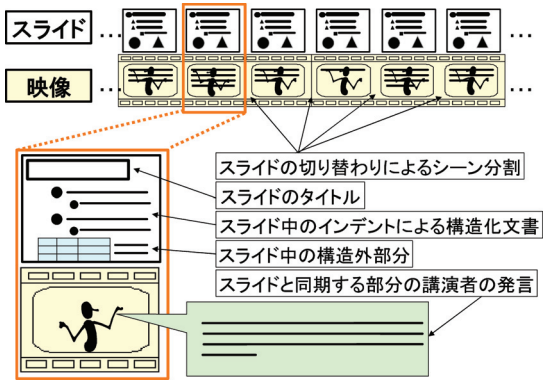


図 2　プレゼンテーションコンテンツの構造
Fig. 2　Metadata of presentation contents.

© 2009 Information Processing Society of Japan

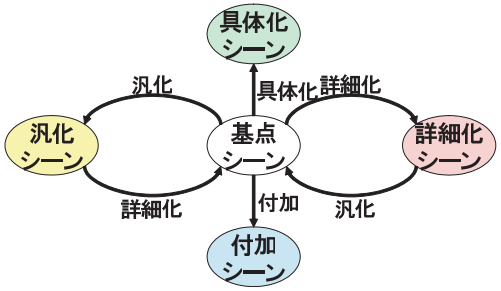


図 3　シーン間の意味的関係
Fig. 3　Relation of semantic scene type.

3.2　コンテンツにおけるシーン間の関係

コンテンツのシーンは，他のシーンと特定の関係を持つ場合があり，本稿では，詳細化，汎化，具体化，付加の 4 種類の関係を定義する．プレゼンテーションでは，単に発表内容を説明する以外にも，聴衆の理解を助けたり興味を引いたりということが行われる．そのため，プレゼンテーションにおける関係には，大きく分けて，あるテーマの内容を説明するための関係，あるテーマを効果的に伝えるための関係の 2 種類があると考えた．

あるテーマの内容を説明するための関係とは，プレゼンテーションを行うあるテーマに対するサブテーマのように，あるテーマを詳細に分割して説明する構造を表す関係である[13]．このとき，あるテーマを説明するシーンと，サブテーマを説明するシーンには詳細化の関係が存在し，逆にサブテーマに対するテーマのシーンには汎化の関係が存在すると考えた．この詳細化および汎化の関係により，あるテーマを詳細に分割した内容を交えて説明する構造として，階層構造が構成される．この階層構造は，あるキーワードに関する詳細化関係もしくは汎化関係を有向辺と見なしたとき，DAG（Directed Acyclic Graph）となる構造である．

あるテーマを効果的に伝えるための関係として，具体例や余談を用いることがある[13]．具体例の場合，特定のシーンに対する具体化という関係を考え，余談の場合，他のシーンに関する余談を述べているため，その対象のシーンに付加されるという関係を考えた．

プレゼンテーションのシーン間の関係を**図 3** に示し，以下に関係の定義を述べる．

詳細化　あるシーンから他のシーンに対して任意のキーワードに関してより詳しい内容を説明しているという関係．用いているスライドがあるキーワードについて特化した内容となり，かつ，そのキーワードに関して講演者が多く言及するという特徴が考えられる．

汎化　あるシーンから他のシーンに対して任意のキーワードに関してより概略の内容を説明しているという関係．用いているスライドがあるキーワードについて一般化した内容となり，かつ，そのキーワードに関して講演者の言及量が減少するという特徴が考えられる．すなわち，詳細化の逆方向の関係が汎化であり，対となる関係といえる．

具体化　あるシーンで説明された任意のキーワードに関して，他のシーンで具体的な内容を説明しているという関係．2 つのシーンで用いているスライドにあるキーワードが出現し，かつ，そのキーワードに関する説明が変化するため講演者の言及量が減少するという特徴が考えられる．なお，具体化はあるテーマを効果的に伝えるためのシーン間の関係であり，あるシーンに対して具体化となるシーンを関係づけると効果的であることしか表さないため，対となる関係は定義しない．

付加　あるシーンで説明された任意のキーワードに関して，他のシーンでもとのシーンと関連性が低いスライドを用いてもとのシーンの内容を説明しているという関係．用いているスライド自体の内容には関係がないため，あるキーワードの記述がスライド中で減少する．一方，口頭での説明はされるため両方のシーンのビデオにあるキーワードが出現するという特徴が考えられる．なお，付加は具体化と同様に，あるシーンに対して付加となるシーンを関係づけると効果的であることしか表さないため，対となる関係は定義しない．

3.3　意味的関係の特性

発表者によってプレゼンテーションの行い方，構成は異なるため，2 つのシーン間の相対的な関係を抽出する．意味的関係の特徴を以下に説明する．

- 同じシーン間でもキーワードごとに異なる関係を持つことが可能である．たとえば，シーン A からシーン B に対してあるキーワード k_1 によって詳細化と判定された場合でも，シーン B からシーン A に対してキーワード k_2 によって詳細化，すなわちシーン A からシーン B に対する汎化と判定されることがありうる．
- 同じシーン間の同一キーワードであっても，複数の関係を持つ場合がある．たとえば，果実というキーワードがスライドのタイトルに含まれ，赤の果物について詳しく説明がされているシーンにおいて，“赤” というキーワードがスライド中で上位の階層に出現し，ビデオ中での発言量が多い場合を想定する．このとき，りんごというキーワードがタイトルに含まれるシーンにおいて，“赤” というキーワードが，下位の階層に出現し，かつビデオ中での発言量が少ないのであれば，りんごというキーワードがタイトルに含まれるシーンは，果実というキーワードがタイトルに含まれるシーンにおける “赤” を一般

© 2009 Information Processing Society of Japan

的に説明しており，かつ “赤” に関して具体的に説明していると見ることも可能である．

- 同じ階層に出現するキーワードは詳細化・汎化の関係にならない．たとえば，「果物その 1」，「果物その 2」，「果物その 3」といったタイトルのシーンがあった場合，“果物” というキーワードに関しては，詳細化も汎化もされない．
- シーン間の関係はプレゼンテーションの構成順序に依存しない．たとえば，発表スライド全体の最初のスライドに，あるキーワードの説明があり，最後のスライドに再度同一キーワードに関する説明があるような場合がある．最初のスライドにおけるあるキーワードに関して最後のスライドが詳細な内容となっていれば，詳細化と判定されることが正しく，一般的な内容となっていれば汎化と判定されることが正しい．

4. シーン間の意味的関係の判定

4.1 具体化判定

意味的関係の判定手法について説明する．スライドでの階層構造の変化やビデオでの発言の割合の変化を用いて判定を行う．まず，具体化の判定について述べ，詳細化および汎化，付加と続けて説明する．判定には，各関係の特徴を考慮した異なる条件式を用いる．

基点シーン i から候補シーン j に対してキーワード k が以下の条件式を満たす場合，2 シーン間に具体化の関係があると定義し，具体化の関係を表す関数 $instance(k,i,j)$ を定義する．

$$slide_num(k,i) > 0 \tag{1}$$

$$slide_num(k,j) > 0 \tag{2}$$

$$K(k,i) = \{m | level(k,i) \leq level(m,i) \text{ and } thread(k,m,i)\} \tag{3}$$

$$\frac{video_num(K(k,i),i)}{video_count(i)} > \frac{video_num(K(k,i),j)}{video_count(j)} \tag{4}$$

$$N(i,j) = \{n \mid slide_num(n,i) = 0 \text{ and } slide_num(n,j) > 0 \text{ and } \frac{video_num(n,i)}{video_count(i)} < \frac{video_num(n,j)}{video_count(j)}\} \tag{5}$$

$$\frac{|N(i,j)|}{slide_keyword(j)} > \alpha \tag{6}$$

条件式 (1)，(2) により，k が基点スライド i と候補スライド j の両方に出現することを判定し，条件式 (4) により基点ビデオ i より候補ビデオ j で k に関するキーワードの割合が減少していることを判定する．

$slide_num(k,i)$ は，k の i 番目のスライドにおける出現数を返す関数である．式 (3) 中の $level(k,i)$ は，i 番目のスライドに出現する k の出現階層を返す関数であり，1 以上の値を返す．k が複数回出現する場合には最も小さい値，すなわち出現する中でより上位の階層を返す．これは，タイトルにキーワード k が出現している場合，本文の箇所で出てくる k はタイトルに出現している k を説明するために出現しているものと考えられるためである．$K(k,i)$ は i 番目のスライドに出現するキーワード k に関連するキーワード集合を返す関数であり，m は k と直接の上下関係にあるインデントに含まれるキーワードである．$thread(k,m,i)$ により，i 番目のスライドにおいて，k と同じ系列のインデントに m が含まれることを判定しており，ビデオ中で入力キーワードに関係あるキーワードの増減を考慮するために用いる．

k に関するキーワードの i 番目のビデオにおける出現の割合は，$video_num(K(k,i),i)/video_count(i)$ により算出する．$video_num(K(k,i),i)$ は，k に関するキーワード集合 $K(k,i)$ におけるそれぞれのキーワードの i 番目のビデオにおける出現数の合計を返す関数である．なお，キーワード集合 $K(k,i)$ ではなく，キーワード k のみを与えることも可能であり，この場合はキーワード k の出現数を返す．$video_count(i)$ は，i 番目のビデオの全キーワードの出現数の対数を返す関数である．

具体的な説明を行っているシーンであるかどうかを判定するために条件式 (6) によって説明に使われている新出語句の割合が閾値 α 以上かどうかを確認する[⋆1]．$slide_keyword(j)$ は j 番目のスライドに含まれる全キーワード種類数を返す関数である．$N(i,j)$ は説明に用いられたと考えられる新出語句の集合である．これは，シーンを変えて具体的に説明する場合，具体的な内容を表すためのキーワードが追加されると考えたためである．

4.2 詳細化・汎化判定

詳細化関係と汎化関係は対になる関係であり，詳細化と判定することができる場合，その逆方向が汎化関係となる．基点シーン i から候補シーン j に対してキーワード k が以下の条件式を満たし，他のシーンおよび他のキーワードに関しての条件に合致した場合，2 シーン間に詳細化の関係があると定義する．以下の条件式により，詳細化の関係の候補であることを表す関数 $detailed_candi(k,i,j)$ を定義する．

$$level(k,i) > level(k,j) \tag{7}$$

$$\frac{video_num(K(k,i),i)}{video_count(i)} < \frac{video_num(K(k,i),j)}{video_count(j)} \tag{8}$$

⋆1 予備実験により閾値は 40%とした．

© 2009 Information Processing Society of Japan

条件式 (7) により，k が基点スライド i より候補スライド j で上位階層に出現していることを判定し，条件式 (8) により基点ビデオ i より候補ビデオ j で k に関するキーワードの割合が増加していること判定する．キーワード k の上位階層での出現を判断基準とするのは，スライドの構造上，上位階層に出現するキーワードはタイトルに近い扱いになると考えられ，スライドがそのキーワードに特化した内容になっていると判断できるためである．

他のキーワードおよび他のシーンに関しての条件として，以下の条件を用いる．

$$\{x|detailed_candi(k,j,x)\} = \phi \tag{9}$$

$$\{k'|(detailed_candi(k',i,j) \text{ or } instance(k',i,j)) \text{ and } k' \neq k\} \neq \phi \tag{10}$$

条件式 (9) は，候補シーン j から，k に関して詳細化となる任意のシーン x が存在しないことを判定するために用いる．条件式 (10) は，基点シーン i から候補シーン j に対して，k 以外のキーワードに関して詳細化または具体化の関係があることを判定するために用いる．条件式 (9) を満たす場合，もしくは満たさないが条件式 (10) を満たす場合に，詳細化関係であると判定する．単一キーワードによる関係判定を行った際，偶然に条件式 (7)，(8) を満たす場合が存在する．段階的に詳細に説明をしている場合，最も詳細な説明をしているシーンであれば，1 つのキーワードだけを詳細に説明しているだけという場合がありうる．そのため，候補シーンがあるキーワードにおいてそれ以上詳細なシーンが存在しないと判定できるのであれば，基点シーンから候補シーンに対して詳細化の関係が存在すると考えた．また，詳細な説明の途中段階にあるシーンであれば，複数のキーワードに関して詳細な説明や具体的な説明が行われており，さらに詳細なシーンで個別のキーワードに関しての詳細説明がされる．そのため，基点シーンから候補シーンに対して他のキーワードでも，詳細化や具体化の関係が判定されるならば，判定中のキーワードも詳細化の関係であると考えた．そのため，これらの条件を単一キーワードによる関係判定に関して補助的に用いる．**図 4** に，基点シーンから候補シーンに対して詳細化と判定される例を示す[⋆1]．

4.3 付加判定

基点シーン i から候補シーン j に対してキーワード k が以下の条件式を満たし，他のキーワードおよび他のシーンに関しての条件を満たす場合，2 シーン間に付加の関係を定義する．以下の条件式により，付加の関係の候補であることを表す関数 $additional_candi(k,i,j)$ を定義する．

⋆1 ここでは k_1 の出現のみを表記しており，シーンには他のキーワードも含まれる．

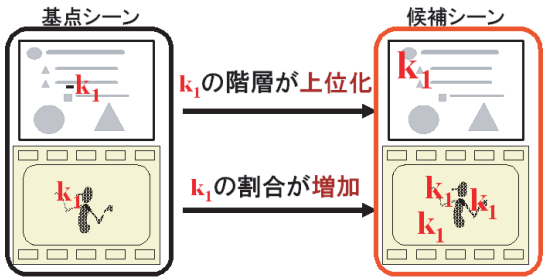


図 4　詳細化判定の例
Fig. 4　Example of detailed relationship.

$$\frac{slide_num(k,i)}{slide_count(i)} > \frac{slide_num(k,j)}{slide_count(j)} \tag{11}$$

$$video_num(k,i) > 0 \tag{12}$$

$$video_num(k,j) > 0 \tag{13}$$

$$level(k,i) < \frac{max_level(i)}{2} \text{ or } level(k,j) > \frac{max_level(j)}{2} \tag{14}$$

条件式 (11) により，基点スライド i より候補スライド j で k の割合が減少していることを判定し，条件式 (12)，(13) により，基点ビデオ i と候補ビデオ j の両方に k が出現することを判定する．さらに，条件式 (14) により k が基点スライド i の中心階層以下から候補スライド j の中心階層以上へ出現階層が変化しないことを判定する．この条件は，指定キーワードがスライド中の内容を表すような階層から，タイトルのような階層へ変化したシーン間では，他の条件を満たしていても，付加的な内容を説明しているとは考えにくいために用いている．

k の i 番目のスライドにおける出現の割合は，$slide_num(k,i)/slide_count(i)$ により算出する．$slide_count(i)$ は，i 番目のスライドのキーワード数の対数を返す関数である．式 (14) 中の $max_level(i)$ は i 番目のシーンにおけるスライドの最深階層を返す関数である．

他のキーワードおよび他のシーンに関する条件は，以下の条件式によって判定する．

$$\{k'|additional_candi(k',i,j) \text{ and } k' \neq k\} \neq \phi \tag{15}$$

$$A(i,j) \not\subset A(\forall x,i) \tag{16}$$

$$A(i,j) = \{a|additional_candi(a,i,j)\} \tag{17}$$

これらの条件式により，基点シーン i から候補シーン j に対して，k 以外のキーワードに

© 2009 Information Processing Society of Japan

関して付加の関係があることと，任意の他のシーン x から基点シーン i に対して付加と判定されたキーワード集合と基点シーン i から候補シーン j に対して付加と判定されたキーワード集合が部分集合とならないことを判定する．また，$A(i,j)$ は，シーン i からシーン j に対して，付加関係の候補と判定されるすべてのキーワード a の集合を表す．

条件式 (15) により，シーン間に他に付加に判定されるキーワードが述べられていることを判定している．これは，あるシーンの補足的な説明を行う場合，その説明のために用いるキーワード数個が同様に付加の関係としてシーン間に存在すると考えられるためである．条件式 (16) は，他のシーンとの間に付加の関係が存在する場合に，それらのシーン間で付加となったキーワードと判定中のシーン間で付加となるキーワードに包含関係が存在しないことを判定している．たとえば，シーン x からシーン i に対して，a，b，c のキーワードによる付加関係があり，シーン i からシーン j に対して，a，b のキーワードによる付加関係がある場合では，キーワード a，b に関して，連続的に付加する説明を行っていることになり，シーン i の内容に対して付加する説明ではないと考えられるためである．

5. 応用例：視聴アプリケーション

5.1 視聴アプリケーションにおける型判定の特性

本稿では，応用アプリケーションとして，シーン抽出および抽出されたシーン群の視聴を行うアプリケーションを生成する．応用アプリケーション生成は，2 つのステップからなる．まず，型判定を行うためのシーン集合の抽出ルールを決定し，次にアプリケーションの目的に応じた型判定ルールを決定する．

視聴アプリケーションにおける，シーン集合の抽出ルールを以下に述べる．

- 任意のシーン中におけるキーワードを入力とする（以降，指定キーワードと記す）．
- 指定キーワードのシーンから，任意のキーワードでいずれかの意味的関係となる他のシーンを抽出する．

視聴アプリケーションのための型判定ルールを用いた視聴順序決定について説明する．抽出したシーン集合に対し適切な視聴順序を決定するための関係として，導入，並列，直列，結論の 4 種類を判定する．これらの型判定による関係は，ある性質のシーン集合の要素である関係と定義する．

導入　特定のシーンに対して汎化関係を持つシーン集合の要素であるという関係であり，特定のシーンを特に導入シーンと呼ぶ．導入シーンは，他のシーンで述べられている内容の概略を説明するシーンであり，視聴アプリケーションでは，導入シーンを他のどのシーンよりも先に視聴するべきシーンとする．

並列　対等な位置づけであるシーン集合の要素であるという関係であり，これらのシーンの順序は可換である．

直列　特定の関係で連なるシーンリストの要素であるという関係であり，これらのシーンの視聴順序は一意に決定する．

結論　特定のシーンに対して詳細化関係を持つシーン集合の要素であるという関係であり，特定のシーンを特に結論シーンと呼ぶ．結論シーンは，複数のシーンの内容を詳細にまとめた内容が述べられるシーンであり，視聴アプリケーションでは，結論シーンを他のどのシーンよりも後に視聴するべきシーンとする．

表 1　型判定による順序決定
Table 1　Order detection rule.

	視聴アプリケーションにおける順序決定ルール
並列	i，j が並列ならば順序関係は不定
直列	i，j が直列ならば $i \prec j$
導入	i が導入シーンならば $i \prec x$
結論	i が結論シーンならば $x \prec i$

視聴順序の決定に関して**表 1** にまとめる．表中の i，j，x は任意のシーンである．すべての型が判定される場合には束構造となるが，抽出したシーン集合にすべての関係が含まれるとは限らないため，これらの型を意味的関係の結果を複合的に解析することでボトムアップに判定を行う．

コンテンツの特徴に対する，型判定について説明する．複数のスライドにわたって階層が同じで 1 つのキーワードについて説明する場合は，そのキーワードで直列と判定されることはなく，そのキーワードについて，同じ種類の関係になるシーンが存在する場合は，並列と判定される．またシーンの順序に依存せず判定を行うため，発表スライド全体の最初のスライドに，あるキーワードの説明があり，最後のスライドに再度同一キーワードに関する説明があるような場合，必ずしも最初のスライドのシーンが導入シーン，最後のスライドのシーンが結論シーンと判定されるとは限らない．

5.2 型判定ルール

導入の関係となるシーン群は，特定のシーンに対して指定キーワードで汎化の関係が成り立つと考えた．以下のいずれかの場合にシーン群の関係を導入と判定し，特定のシーン s を導入シーンと判定する．

© 2009 Information Processing Society of Japan

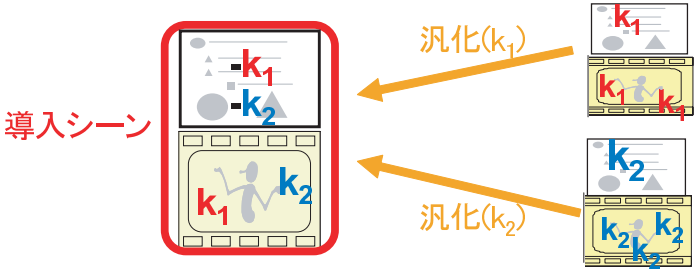


図 5　導入判定の例
Fig. 5　Example of introduction scene.

- 複数のシーンからシーン s に対して，指定キーワード k による汎化の関係が成り立つ．
- 任意のシーンから指定キーワード k により汎化の関係が成り立つ候補シーン群から，シーン s に対して任意のキーワードによる汎化の関係が成り立つ．

導入シーンは，導入のシーン群を構成する際に，汎化関係で有向グラフを作成した際に，終点となるシーンである．**図 5** に，シーン群が導入であり，太線で囲まれたシーンが導入シーンと判定される例を示す．ここではキーワード k_1，k_2 の出現のみを表記しており，シーンには他にも多数のキーワードが含まれる．

並列の関係を持つシーン群は，特定のシーンに対して任意のキーワードによって同一の関係が成り立つと考える．以下の手順により，並列の判定を行う．

(1)　共通のシーンに対して任意のキーワードによる関係が同一となるシーン群を抽出．
(2)　共通のシーンを除いたシーン群中におけるシーン間の関係を判定し，詳細化もしくは汎化の関係となるキーワードがあれば並列としない．

以上により，並列の関係を持つシーン群を判定する．シーン間に詳細化，汎化の関係がある 2 シーンを並列としないのは，内容の詳細度に差があり，プレゼンテーション内で対等な位置づけであるシーンという定義を満たさないためである．

直列の関係を持つシーン群は，特定の関係により連なったシーン群に相当すると考える．そこで，任意のキーワードに関する関係を用いて有向グラフを作成し，各シーンを始点として関係をたどることにより，連なったシーン群を直列と判定する．関係をたどる手順を以下に示す．

(1)　始点となるシーンから汎化，詳細化，または具体化の関係をたどる．
(2)　終了条件に至るまで，次のシーンからさらに関係をたどる．
- 始点から汎化の関係でたどった場合，汎化，付加，または具体化の関係をたどる．
- 始点から詳細化の関係でたどった場合，詳細化，付加，または具体化の関係をたどる．

(3)　以下の条件を満たせば終了する．
- 具体化の関係をたどる．
- たどることができる関係が存在しない．
- 同シーン間の同関係を 2 度たどる．

以上により，関係の連なったシーン群を直列と判定する．

結論の関係を持つシーン群は，特定のシーンに対して指定キーワードで詳細化の関係が成り立つと考えた．以下のいずれかの場合にシーン群の関係を結論と判定し，特定のシーン s を結論シーンと判定する．

- 複数のシーンからシーン s に対して，指定キーワード k による詳細化の関係が成り立つ．
- 任意のシーンに対して指定キーワード k により詳細化の関係が成り立つ候補シーン群から，シーン s に対して任意のキーワードによる詳細化の関係が成り立つ．

すなわち，指定キーワードに関して，詳細化関係で有向グラフを作成した際に，終点となるシーンが結論シーンであり，結論シーンに上記の条件で接続しているシーン群が結論となる．

6. プロトタイプシステム

6.1 システム構成

プロトタイプシステムは Visual Studio 2008 の C# によって構築した．スライド，ビデオからの単語抽出は SlothLib[14),15)] より呼び出した茶筌[16),17)] による形態素解析を用いて行った．

図 6 に構築したプロトタイプシステムの画面イメージを示す．入力画面では，スライド画像の一覧と，画像の横にシーンで出現するキーワードリストを表示する．ユーザはこのキーワードリストから，指定キーワードを選択することができる．出力画面では，指定したシーンと他のシーンの関係をグラフで見ることができる．再生ボタンがクリックされた場合，別ウインドウで抽出区間を再生する．ユーザは出力画面に提示されるシーン間の関係を用いることで，プレゼンテーション全体における視聴中のシーンの位置づけを把握することが可能であり，抽出区間の再生中に該当シーン間の関係を提示することで，シーンの切り替わり時に注目するべき関係を得ることが可能となる．

© 2009 Information Processing Society of Japan

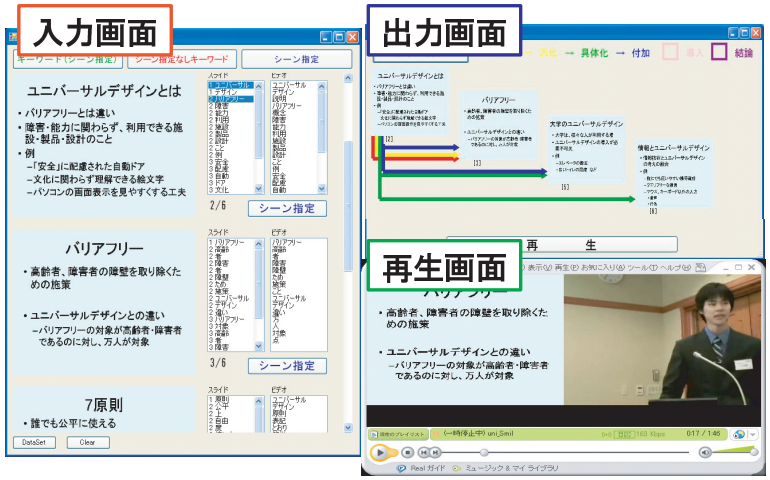


図 6　プロトタイプシステムの画面イメージ
Fig. 6　Screen image of the prototype system.

6.2　予備実験：意味的関係の網羅度

本稿で定義しているシーン間の関係の網羅度を確認するために予備実験を行った．網羅度を算出することで，コンテンツ内に含まれる関係の多くが定義している 4 種類の意味的関係であることを確認する．プレゼンテーション内の 2 シーンと，2 シーンに出現するキーワードを被験者に提示することで，自由記述によるキーワードに関するシーンの関係を得た．被験者は 5 名であり，2 種類のコンテンツからランダムにサンプリングした 2 シーンを 60 セット，2 シーンに含まれる各 5 キーワードに関して，計 300 のシーン間の関係を判定した．用いたコンテンツは，日本データベース学会の DBSJ Archives[18] で公開されている DEWS2004 で行われた「モバイル環境における端末の位置情報に基づく P2P ネットワークの提案と評価」というコンテンツ，および DEWS2006 で行われた「選言的木パターンマイニング」というコンテンツである[⋆1]．自由記述により何らかの関係があると 3 名以上が記述したシーンを全体集合とし，自由記述による回答に関して，定義している意味的関係もしくはその他のいずれに相当するのかをヒアリングした結果を部分集合として，意味的関係の網羅度を算出した．

⋆1 6.3，6.4，6.5 節で行っている評価実験も同じコンテンツを用いている．

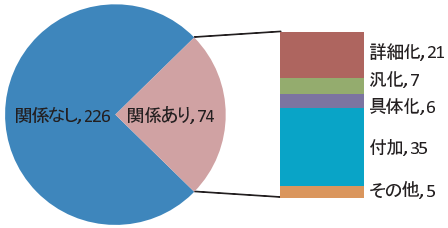


図 7　意味的関係の網羅度
Fig. 7　Coverage of semantic relation.

結果を**図 7** に示す．その他という回答になったものは 5 組のシーンであり，網羅度は 93.2%（69/74）という結果となった．他の関係として回答されたものは，比較，並列，連続といったものであった．その他の関係であっても，たとえば視聴アプリケーションにおいて定義している並列関係のように，これらの関係は型判定によって定義可能であると考えられる．すなわち，定義している意味的関係でほぼ網羅できるものと考えられる．

6.3　評価実験 1：意味的関係の妥当性

6.3.1　実験環境と実験結果

本手法で定義している意味的関係の妥当性を評価するために実験を行った．被験者はシーン間の関係が，詳細化，汎化，具体化，付加，その他のいずれに該当するかを判定した．被験者は 5 名であり，実験データとして，2 種類のコンテンツからランダムにサンプリングした 2 シーンを 60 セット[⋆2]，各基点シーンに含まれる 5 つのキーワードを用い，それらに関して，計 300 のシーン間の関係を判定した．用いたシーン組とキーワードの一部を**表 2**，**表 3** に示す．それぞれの関係は独立に判定されるため，詳細化関係と具体化関係など，1 つのキーワードに対し複数の関係がシーン間に付与される場合もある．実際にユーザが利用する場合を想定するために，基点シーンを現在視聴中のシーンと仮定し，シーン中に含まれるキーワードから選択される可能性が高いと判断した 5 つのキーワードを抽出している．そのため，実験に用いたシーン組中には，実験に用いたキーワードが片方にしか含まれていないなど，システムが適切に判定することができないと考えられるシーン組が多数含まれている．被験者が判定した結果を正解として，システムの判定結果を用いた適合率・再現率・F

⋆2 本実験では，すべての組合せからランダムに抽出したため，シーン A に対するシーン B と，シーン B に対するシーン A は別の組合せとして抽出される．

© 2009 Information Processing Society of Japan

表 2　実験データ：モバイル環境における端末の位置情報に基づく P2P ネットワークの提案と評価
Table 2　Experimental data: LL-Net: location-based logical P2P network for mobile environment.

基点シーン	2	3	3	4	4	5	5	6
候補シーン	10	4	15	6	14	9	11	3
キーワード	モバイル	コンテキスト	ネットワーク	ネットワーク	管理	位置	プロトコル	階層
	情報	ネットワーク	場所	検索	分散	ホップ	端末	提案
	環境	営業	趣向	自律	情報	エリア	提案	端末
	利用	混雑	故障	技術	環境	問題	故障	エリア
	街角	位置	状況	実現	位置	学会	転送	評価

表 3　実験データ：選言的木パターンマイニング
Table 3　Experimental data: mining disjunctive tree patterns.

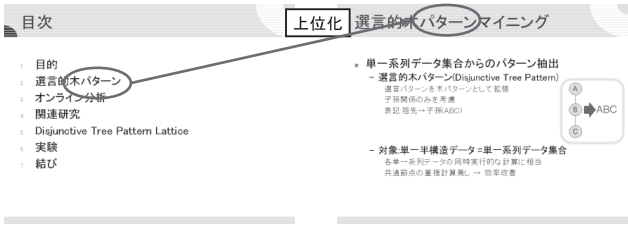
基点シーン	2	3	3	4	5	6	7	8
候補シーン	7	15	19	3	10	20	16	11
キーワード	目的	データ	系列	レベル	選言	系列	拡張	特徴
	選言	パターン	集合	ラベル	木	カウント	選言	テキスト
	パターン	有向グラフ	研究	節点	パターン	単調	パターン	タグ
	オンライン	ラベル	構造	系列	概念	データ	マイニング	パターン
	実験	抽出	選言	集合	辞書	実験	子孫	従来

表 4　判定結果の分類
Table 4　Experimental result.

		システム判定結果				
		詳細化	汎化	具体化	付加	その他
正解	詳細化	**18**	7	19	13	95
	汎化	4	**19**	10	15	23
	具体化	2	8	**12**	14	20
	付加	15	26	24	**54**	159
	その他	11	25	25	54	948

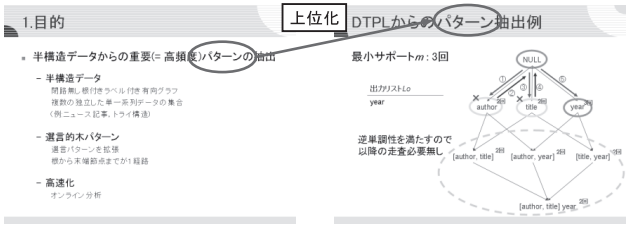
値により評価を行った．

判定結果の分類を**表 4** に示す．表 4 は，縦にシステム判定結果，横に被験者判定結果による正解を示す．表中の下線が引かれた数値は，システムの判定結果と正解が一致した解の数である．また，システムの判定結果におけるその他とは，判定不能であったデータである．詳細化に関して適切に判定されているシーン組の例を**図 8** と，適切に判定することができていないシーン組の例を**図 9** に示す．意味的関係の判定結果を**表 5** に示す．なお，システムが複数の種類の関係を判定するのに対し，被験者は 1 つの関係を判定しているため，表 4 に示した被験者判定結果である正解には重複があり，再現率の算出には重複を除いたものを用いている．



パターン発言数：1
全キーワード発言数：77
<
パターン発言数：4
全キーワード発言数：75

図 8　詳細化判定の成功例
Fig. 8　Adequate detecting the detail.



パターン発言数：3
全キーワード発言数：38
>
パターン発言数：2
全キーワード発言数：70

図 9　詳細化判定の失敗例
Fig. 9　Inadequate detecting the detail.

表 5　意味的関係の評価
Table 5　Result of detected relation type using related keywords.

	詳細化	汎化	具体化	付加	全体
適合率	36.0% (18/50)	22.4% (19/85)	13.3% (12/90)	36.0% (54/150)	27.5% (103/375)
再現率	12.8% (18/141)	40.4% (19/47)	29.3% (12/41)	22.2% (54/243)	21.8% (103/472)
F 値	0.19	0.29	0.18	0.27	0.24

6.3.2　意味的関係判定の考察

この評価実験により，シーン間のスライド階層構造を用いる条件の妥当性を確認した．以下に関係別の考察を述べる．

© 2009 Information Processing Society of Japan

詳細化・汎化　本手法により，スライドに準じてあるキーワードの説明を詳細化させているような場合についての判定を行うことが可能であると考えられる．たとえば，図 8 の場合のように，目次のようなシーンからその内容を説明するシーンに対しては，適切に判定することが可能である．一方，図 9 のように，シーンとしては “パターン” について詳細な説明が行われているにもかかわらず，発言量として増加しないため適切に判定することができていない．

具体化　本手法によって扱うことが可能な具体化は，スライド中に「説明対象のトピックを表現するキーワード」が記載されているような場合であると考えられる．たとえば，“ネットワーク” というキーワードに関して，“ネットワーク” というキーワードを用いずに “ピア” の説明をしているシーンがある．このような具体化に関しては，単一キーワードの変化によってシーン間の関係を判定する本手法では扱うことができない．しかしながら，“ネットワーク” の例であれば，“ピア” を説明するスライド中に，“ネットワーク” のトピックであることを示すためにキーワードが記載されていれば判定可能であると考えられる．

付加　付加の判定では，多くの正解がシステムにおいていずれの意味的関係にも判定されず，無関係となる傾向が見られた．これは，スライドにおける説明の度合いとキーワードの出現割合が一致しないことが一因と考えられる．

評価実験に用いたコンテンツは研究発表のコンテンツのみを用いている．研究発表のコンテンツの傾向として考えられる特徴を以下に述べる．

- 詳細化が他の関係に比べ，適合率が比較的高い値となったのは，実験に用いたコンテンツが研究発表であり，ある事柄に対して詳細に説明を行うというシーンが多数あるためと考えられる．
- 汎化関係というラベルを付与して提示することで理解の促進が可能であると考えられる結果を得た．研究発表のコンテンツであり，内容をある程度理解した被験者でなければ，正解を正しく判定することが困難であると考えられる．そのため，同じシーンの組合せに対して，ある被験者は「一般的に説明している」と回答したが，別の被験者は「関係がある別のテーマ」と回答するなど，汎化と付加に回答が別れるケースが多く見られた．これは，一般的な概念であると理解できる場合には汎化と解釈することが可能であるが，理解できなかった場合には関係することまでは分かるがどのような関係か理解できないようなシーンであり，システムによって汎化関係と示すことが有効なシーンであると考えられる．
- 具体化は被験者の解に比べ，システムの解の数が非常に多い結果となった．研究発表のコンテンツでは，具体例に関する説明が内容を説明したシーン内で行われることが多く，別シーンで具体例のみの説明を行うというケースが少ないということが考えられる．
- 付加の被験者解が多数存在したのは，基本的に研究発表では無関係なことを述べるようなシーンはないため，付加することで効果的に理解可能なシーンとしては，多くのシーンが該当するためと考えられる．

評価実験では判定精度が低い結果となった．そのため，以下のような判定手法の改良が必要であると考えられる．

- スライド中におけるキーワードの階層関係を重視した判定とすることで判定精度の改善を行うことができると考えられるが，提案手法で行っているスライド間のキーワードの上位階層化に関して，スライド中の絶対的な階層による判定のみではなく，スライド中に出現している他のキーワードから見た相対的な階層による判定を組み合わせて行う方法などを検討する必要があると考えられる．
- シーンの単位に関して，1 つのシーン内で内容と具体例を述べるといったことが起こるため，適切なシーンの区切りに関する検討が必要であること，また，判定条件を絞るために，特定の関係でのみ現れるキーワード（“たとえば” など）を用いることの検討が必要であることを確認した．

6.4　評価実験 2：型判定の妥当性

6.4.1　直列判定の妥当性

本手法で定義している直列判定の妥当性を評価するために実験を行った．直列は順序を持ったシーン群である．そこで，本稿で定義する直列判定により，順序が妥当であるシーン群が取得できるかどうかを確認する．プレゼンテーション内のシーンを並べて被験者に提示し，シーン群の順序の妥当性の評価を行った．被験者に提示するシーン数を，2〜4 シーンと変化させ，シーン数ごとに評価を行った．被験者は 4 名であり，6.3 節で得られた正解のみを用いたデータセットによる実験により，順序が妥当である組合せを判定した．被験者によって順序が妥当であるという回答が得られたシーン群を正解として，システムの判定結果を用いた適合率により評価を行った．結果を**表 6** に示し，考察を以下に述べる．

実験の結果，シーン数が少ない場合であれば，被験者によって「連続して視聴することが妥当である」という回答を多く得ることができ，直列関係の判定は適切に行うことが可能であると考えられる．直列となるシーン数が多くなる場合であれば，シーン群を連続して視聴したときに，最初のシーンと最後のシーンで説明しているトピックが異なるシーン群が抽出

© 2009 Information Processing Society of Japan

表 6　直列判定の評価
Table 6　Result of detected inline scene.

	2 シーン	3 シーン	4 シーン	全体
適合率	83.9% (26/31)	56.0% (14/25)	21.4% (3/14)	61.4% (43/70)

されやすくなるため，精度が著しく下がる結果となった[⋆1]．この問題に対しては，抽出されたシーン数に応じて，あらかじめ適切な長さを定めておくことで対処することが考えられる．また，他の誤判定としては，システムが，汎化を基準に関係をたどることで直列と判定したシーン群に対し，被験者が順序が逆であると解答する場合が見られた．

6.4.2　並列判定の妥当性

本手法で定義している並列判定の妥当性を評価するために実験を行った．被験者は 4 名であり，6.3 節で得られた正解のみを用いたデータセットに関して行った．被験者が抽出されたシーンが対等であると判定した結果を正解として，システムの判定結果を用いた適合率により評価を行った．4 個のシーン群が抽出され，そのうち 2 個に関して，「対等なシーンである」という回答が得られたため，適合率が 50.0%という結果となった．考察を以下に述べる．

被験者によって「2 つのシーンは対等の内容を述べている」という回答を得ることができ，並列の型判定は適切に行える可能性があることを確認した．並列判定の手順 1 によって，10 個のシーン群が抽出され，手順 2 により 4 個に絞り込むことができた．今回の実験では，除外された 6 個のシーンに正解は含まれない結果となり，並列の型に関係ないシーンの除外が可能であると考えられる．誤りとなった原因としては，関係を判定したキーワード間にまったく関係がない場合に，意味的関係の種類が同じであっても，抽出されたシーンが無関係であり，対等ではないことがあげられる．

並列判定においては，すべての意味的関係の判定精度に影響を受ける．適切な並列関係を得るためには，最低 2 つ以上のシーンと同じキーワードで同じ関係となる必要がある．6 シーン抽出された場合を想定すると，何らかの関係で，33.3%程度の再現率があれば，適切な並列関係の判定を行うことが可能であると考えられる[⋆2]．しかし，どれか 1 つの関係でも適合率が低い場合には，多数の並列関係が抽出される結果となる．コンテンツによって，あらかじめ特定の関係が抽出されにくいことが想定される場合，アプリケーション作成時にその関係をたどらないように作成することで対処することが可能であると考えられる．たとえば，今回用いている研究発表のコンテンツであれば，本研究で想定している別のスライドを用いて具体例のみを説明するシーンが出現しにくいことが想定されるため，あらかじめ並列判定に具体化を含めないといった対処が考えられる．

6.4.3　導入・結論判定の妥当性

本手法で定義している導入・結論判定の妥当性を評価するために実験を行った．6.3 節で得られた正解のみを用いたデータセットによる実験を行った．このデータセットから抽出されたシーン群は導入が 0 個，結論が 2 個であった．導入が抽出されなかった原因としては，複数のシーンから 1 つのシーンに対して，同じキーワードで汎化関係となるという条件が必要であるのに対し，用いたデータセットにおいて同じキーワードによる汎化関係が存在しなかったためである．そのため，結論判定のみ実験を行い，その結果から，導入・結論判定の考察を行う．被験者が抽出されたシーン群から詳細にまとめているシーンを選択し，そのシーンがシステムにより結論シーンと判定されているかどうかを評価した．その結果，抽出された 2 個のシーン群のどちらも，被験者の解とシステムの解が一致したため，適合率は 100%となった．考察を以下に述べる．

結論判定に関しては，被験者によって「最も詳細にまとめたシーンである」という回答を得ることができ，結論関係の判定は適切に行える可能性があることを確認した．一方，システムによって導入の関係となるシーン群を得ることができなかった．しかしながら，結論は詳細化，導入は汎化によって判定される関係であり，逆方向の判定を行っていることから，意味的関係判定が適切に行われることで，導入関係も抽出可能であると考えられる．

導入判定においては，指定キーワードに関して，いくつかの汎化関係が判定される必要がある．正しい導入を得るためには，最低 2 つ以上のシーンから適切な汎化の関係を必要とする．あるシーンのキーワードを指定した際に，6 シーン抽出された場合を想定すると，33.3%程度の再現率で汎化を判定することができれば，適切な導入関係と導入シーンの判定を行うことが可能である[⋆3]．このように判定漏れに関しては，対処可能な精度であると考えられるが，誤判定により，判定数が多い場合には，多数の導入関係が抽出される結果になると考えられる．結論判定に関しては，導入と同様の理由で，詳細化判定に影響を受ける．今

⋆1 ランダムにサンプリングしたデータを用いた実験においても同様の傾向が見られた．なお，この場合の適合率は 36.2%，再現率は 74.3%であった．
⋆2 正解のみではなくすべてのデータを用いて実験を行ったところ，適合率が 5.2%，再現率が 63.2%であった．
⋆3 正解データではなく，指定したキーワードによって抽出した出力では，導入の適合率が 20.5%，再現率が 60.3%であった．また，結論は 18.6%，57.1%であった．

© 2009 Information Processing Society of Japan

表 7　提案手法による出力シーンと判定に用いたキーワード
Table 7　The output of our proposed method.

	3	7	6（基点）	5	13	14	9	10	15	16	12（結論）
タイトル	研究背景（2/3）	エリアの階層化	研究目的	端末の位置情報に基づく P2P ネットワーク（LL-Net）	比較対象との比較評価（1/2）	比較対象との比較評価（2/2）	ネットワーク参加プロトコル	エリア間移動プロトコル	故障を考慮した評価（1/2）	故障を考慮した評価（2/2）	シミュレーション評価
詳細化		エリア，階層，化		端末，位置，ネットワーク			ネットワーク	エリア	評価		シミュレーション，評価
汎化		ネットワーク		提案，評価	階層						
具体化		ネットワーク						情報		評価	
付加	情報	提案			ネットワーク，階層，化，提案	ネットワーク					ネットワーク，エリア

表 8　tf-idf 法による出力シーンと判定に用いたキーワード
Table 8　The output of tf-idf method.

	4	5	6	7	17	12
タイトル	研究背景（3/3）	端末の位置情報に基づく P2P ネットワーク（LL-Net）	研究目的	エリアの階層化	まとめ	シミュレーション評価
キーワード	キー	的	シミュレーション	目的	シミュレーション	シミュレーション
tf-idf 値	0.5	0.5	0.67	1.0	1.0	2.67

回行った実験の再現率では，20 シーンほど抽出された場合に適切な結論関係と結論シーンを得ることができると考えられる．

6.5　評価実験 3：関係提示の有効性

6.5.1　比 較 実 験

従来手法として映像のメタデータ抽出を行う informedia[19] における tf-idf 法による特徴語によるシーン検索，および入力キーワードに関してそのキーワードの tf-idf 値順に並べたシーン提示との差異を考察する．研究発表映像では，映像特徴によるシーン検出がほとんど行われないため，シーンの区切りは本手法同様にスライド単位で行い，シーンを単位とした df 値を計算している．比較に用いた「モバイル環境における端末の位置情報に基づく P2P ネットワークの提案と評価」のコンテンツの出力シーンとその判定に用いられたキーワードを**表 7**，**表 8** に示す．提案手法による出力は，シーン 6 中のキーワード “シミュレーション” を指定した場合の出力結果であり，シーン 6 に関してすべてのキーワードによって関係が判定されたシーンのうち，指定キーワードを用いて型判定を行い，シーンの順序を決定したものである．tf-idf 法による出力は，シーン 6 に含まれるキーワードのうち，tf-idf 値上位 3 語と指定キーワード “シミュレーション” による検索結果を該当シーンにおける tf-idf 値順に並べたものである．提案手法による出力画面としては，出力シーンのスライドが出力画面に並び，それらのシーン間の意味的関係が色別に矢印によって表示される．また，型判定によって得られた関係は，スライド集合を線で囲み，関係の名称を併記することで表現している．再生ボタンを押すことで，シーンを連続的に再生する．tf-idf 法による出力画面は，出力シーンのスライドを出力画面に並べ，各シーンの抽出に用いた特徴語を併記した矢印でシーン間をつないで提示したものを想定している．同様に再生ボタンによりシーンを連続的に再生する．被験者は 5 名であり，出力された結果により，指定したキーワードに関して理解に有効である，もしくは有効ではない理由を自由記述で回答した．

6.5.2　実験結果と考察

tf-idf 法による出力では，ユーザが視聴しているシーンに対して理解の助けになるような出力を得られにくい結果となった．シーン 6「研究目的」というシーンでは，“キー”，“的”，“目的” というキーワードが特徴的な語句として抽出された．これらのキーワードは，「研究目的」というタイトルのシーンにおいて，「位置をキーとする」や「階層的に考慮」といった，このシーンの特徴を表しているキーワードであると考えられ，これらのキーワードは，シーン自体の内容の把握には有用であると考えられる．被験者の回答では，キーワードを提示することでシーン中の注目するべき個所が明確になり理解に有効であるという意見が得られた．一方，指定キーワード “シミュレーション” による検索では，シーン 12「シミュレーション評価」のように有用なシーンが検索される一方，シーン 17「まとめ」のように，視聴シーン 6 と同程度の内容しか述べていないシーンも検索されている．tf-idf 値順の出力では，特にシーン間に関係がなく，どのキーワードに着目して視聴すればよいか分からない出力となることが考えられる．被験者の回答では，シミュレーションに関して具体的に説明するシーンが抽出されていないため，深い理解にならなかったという意見が得られた．

© 2009 Information Processing Society of Japan

提案手法による出力では，ユーザが視聴しているシーンに対して理解の助けになると考えられる出力が得られた．直列の型判定により，シーン 7，6，5，13，14 という順序を決定した．なお，導入，並列の型判定に関しては，今回の結果では出力が得られなかった．被験者の回答では，視聴中のシーンに関して，詳細になっている最中である（シーン 7，6，9，10，12）や具体的な情報がある（シーン 7，16）という提示がされることで，説明のつながりを把握することができ有効であるという意見が得られた．しかし，関係のないシーン（シーン 3，13，14）が多数抽出されているため有効ではないという意見が得られた．また，結論の型判定により，シーン 6，12，17 が “シミュレーション” に関して結論となるシーン群でありシーン 12 が最も詳細にまとめている結論シーンであると判定した．なお，シーン 17 はシーン 6 に対して意味的関係が存在しないため，型判定にのみ使われ出力しない．このことにより，ユーザは，説明内容に差がないシーン 17 を省いて，シーン 12 を見ることが可能であり，また，シーン 12 がシミュレーションに関して最も詳細なシーンであるということを知ることが可能となる．被験者の回答では，シミュレーションに関して，一番まとまっているスライドが直感的に分かるのは理解しやすいという意見が得られた．

tf-idf 法による特徴語の提示や，tf-idf 値順の視聴は，シーン単位の内容把握には有効であると考えられるが，シーン間のつながりが不明確であるため，抽出したシーン集合を視聴するのには不向きであると考えられる．提案手法による関係提示は，抽出されたシーン集合に対する内容把握として有効であると考えられる．たとえば，入力したキーワードに関して最もまとまっているシーンを発見するような場合や，入力キーワードに関してプレゼンテーション全体で述べられている内容を理解する場合にはシーン抽出および関係の明示により本手法が有効に作用するものと考えられる．今回の実験では，意味的関係の判定精度が不十分であり，不要なシーンが多数抽出された結果となっている．今後の課題としては，意味的関係の判定精度を改善することで，型判定の定量的な評価を行うことがあげられる．

7. おわりに

本稿では，コンテンツにおけるシーン間の構造的な特徴を分析することで意味的関係を抽出する方式について提案を行い，その応用例を示した．本手法によって，シーンを理解するために必要な他のシーンを判定し，意味的につながりのあるシーンを抽出することが可能となる．また，意味的関係を用いたアプリケーションを作成することによりあるシーン中における指定されたキーワードに応じたシーン群を抽出し，それらの関係を提示することが可能となる．

評価実験の結果，意味的関係判定アルゴリズムが適応できるコンテンツに関して考察を行った．また，応用アプリケーションである視聴アプリケーションを用いて評価を行い，関係抽出を用いたアプリケーションが有効である可能性を確認した．各判定アルゴリズムに関しては十分な精度が得られていないが，改良していくことで，関係を指定したプレゼンテーション検索アプリケーションなど，他のアプリケーションへの応用が可能であると考えられる．プレゼンテーションでは，図を中心に説明する場合や，階層構造をほとんど用いないで発表資料を作成する場合なども考えられるため，スライドにおける階層変化やビデオにおける出現の割合だけでなく，スライドにおける字の色や配置，ビデオにおける音量の変化などを考慮しシーン間の変化を解析していく必要があると考えられる．また，本稿ではキーワードごとに判定を行っているが，プレゼンテーションの内容から適切にキーワードの組合せ判定を行う必要がある．さらに，単一のコンテンツを対象とするのではなく，複数コンテンツを対象とした判定方式への拡張が，今後の課題である．

謝辞　本研究の一部は，平成 21 年度科研費基盤研究（B）(2)「ユーザの潜在的意図を用いたレス・コンシャス情報検索基盤の構築」（課題番号：20300039）によるものです．ここに記して謝意を表します．

参 考 文 献

1) Ricoh Corporation. MPmeister, http://www.ricoh.co.jp/mpmeister/
2) He, L., Sanocki, E., Gupta, A. and Grudin, J.: Auto-Summarization of Audio-Video Presentation, *Proc. 7th ACM international conference on Multimedia* (*Part 1*), pp.489–498 (1999).
3) 吹野直紀，馬　強，角谷和俊，田中克己：ニュース記事を利用したサッカー要約映像の生成，電子情報通信学会第 14 回データ工学ワークショップ（DEWS2003）論文集，pp.8–P–03 (2003).
4) Smith, M.A. and Kanade, T.: Video Skimming and Characterization through the Combination of Image and Language Understanding Techniques, *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (*CVPR 07*) (1997).
5) 小林隆志，村木太一，直井　聡，横田治夫：統合プレゼンテーションコンテンツ蓄積検索システムの試作，電子情報通信学会論文誌，Vol.J88-D-I, No.3, pp.715–726 (2005).
6) 村木太一，吉田　誠，小林隆志，直井　聡，横田治夫：メタデータによる講演資料と動画の統合と検索，データベースと Web 情報システムに関するシンポジウム（DBWeb2002）論文集，pp.97–104 (2002).
7) Nakano, W., Kobayashi, T., Katsuyama, Y., Naoi, S. and Yokota, H.: Treatment of

© 2009 Information Processing Society of Japan

Laser Pointer and Speech Information in Lecture Scene Retrieval, *Proc. 8th IEEE Intl Symp. on Multimedia* (*ISM2006*), pp.927–932 (2006).
8) Pradhan, S., Tajima, K. and Tanaka, K.: A Query Model for Retrieving Relevant Intervals within a Video Stream, *Proc. 6th IEEE Int'l Conference on Multimedia Computing and Systems* (*ICMCS 99*), Vol.2, pp.788–792 (1999).
9) 戈　指夷，角谷和俊：遠隔会議システムにおける資料操作ログに基づくアーカイブコンテンツ作成支援方式，電子情報通信学会第 17 回データ工学ワークショップ（DEWS2006）論文集，pp.6C–i5 (2006).
10) 北山大輔，大谷亜希子，平元綾子，角谷和俊：プレゼンテーションコンテンツのための意味的シーンタイプに基づく視聴区間検索，データベースと Web 情報システムに関するシンポジウム（DBWeb2007），pp.1C–2 (2007).
11) 小山　聡，田中克己：質問の階層的構造化を利用した Web 検索手法の提案，日本データベース学会 Letters（DBSJ Letters），Vol.1, No.1, pp.63–66 (2002).
12) Yumoto, T. and Tanaka, K.: Page Sets as Web Search Answers, *Proc. 9th International Conference on Asian Digital Libraries* (*ICADL2006*), pp.244–253 (2006).
13) 八幡紕芦史：パーフェクトプレゼンテーション，生産性出版 (1995).
14) SlothLib. http://www.dl.kuis.kyoto-u.ac.jp/SlothLibWiki/
15) 大島裕明，中村聡史，田中克己：SlothLib：Web サーチ研究のためのプログラミングライブラリ，日本データベース学会 Letters，Vol.6, No.1, pp.113–116 (2007).
16) Japanese Morphological Analysis System ChaSen. http://chasen.naist.jp/hiki/ChaSen/
17) Asahara, M. and Matsumoto, Y.: Extended Models and Tools for High-performance Part-of-Speech Tagger, *Proc. 18th International Conference on Computational Linguistics* (*COLING 2000*), pp.21–27 (2000).
18) DBSJ Archives. http://www.dbsj.org/Japanese/Archives/archivesIndex.html
19) 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.46–52 (1996).

(平成 20 年 12 月 20 日受付)
(平成 21 年 4 月 7 日採録)

（担当編集委員　山根 康男）



北山　大輔（学生会員）

2005 年姫路工業大学環境人間学部環境人間学科卒業．2007 年兵庫県立大学大学院環境人間学研究科博士前期課程修了．現在，兵庫県立大学大学院環境人間学研究科博士後期課程に在籍．日本学術振興会特別研究員 DC．映像データベース，マルチメディアデータベースに興味を持つ．日本データベース学会学生会員．



大谷亜希子

2008 年兵庫県立大学環境人間学部環境人間学科卒業．同年 NTT データネッツ株式会社入社，現在に至る．



角谷　和俊（正会員）

1988 年神戸大学大学院工学研究科修士課程修了．同年松下電器産業（株）入社．ソフトウェア開発環境，マルチメディアデータベース，データ放送の研究開発に従事．1998 年神戸大学大学院自然科学研究科博士後期課程（情報メディア科学専攻）修了．1999 年神戸大学都市安全研究センター都市情報システム研究分野講師，2000 年同助教授．2001 年京都大学大学院情報学研究科社会情報学専攻助教授．2004 年兵庫県立大学環境人間学部環境人間学科教授，現在に至る．博士（工学）．IEEE Computer Society，ACM，電子情報通信学会，日本データベース学会等各会員．

© 2009 Information Processing Society of Japan