

スライドシーンに基づくスライドリポジトリ構築手法の提案

土井 達也[†] 清水 堅^{††} 伊藤 太樹^{††} 柿元 宏晃^{††}

白松 俊^{††} 大園 忠親^{††} 新谷 虎松^{††}

名古屋工業大学工学部情報工学科[†] 名古屋工業大学大学院工学研究科情報工学専攻^{††}

1 はじめに

プレゼンテーション資料の作成において、スライドの再利用が多く行われている。本稿ではプレゼンテーション用ファイルをプレゼンテーション資料と呼ぶ。スライドとはプレゼンテーション時に表示される 1 画面のシートのことであり、プレゼンテーション資料は複数のスライドから構成される。スライドの再利用を行うためには、保存されているプレゼンテーション資料から必要なスライドを検索し発見する必要がある。現在主流の検索システムは、キーワードの出現頻度に基づいているものが多い。そのため、偏った分野のプレゼンテーション資料が存在する環境において、研究発表や講義などの作成目的に適したスライドの発見は困難となる。作成目的によってプレゼンテーション資料は異なるスライド間構造をもつ。スライド間構造とはスライドの前後関係のことである。

本研究ではプレゼンテーション資料のスライド間構造に着目し、スライドリポジトリを構築する。スライド間構造を考慮することにより、プレゼンテーション資料の作成目的に適したスライド作成支援が行うことができる。構築したスライドリポジトリを利用することでスライド作成を支援する [2]。本稿ではスライド作成支援を行うためのスライドリポジトリの構築手法について述べる。

2 スライド間構造に基づくスライド作成支援

プレゼンテーション資料を作成するにあたって、過去に作成したスライドを再利用することが多い。よって再利用するスライドの検索と発見を支援することでスライド作成支援につながると考えられる。スライドは文字サイズやインデントなどの視覚的な構造を持つ。よってスライドを有効利用するために計算機で処理しやすい形式にスライドを変換することでプレゼンテーション資料にまとめられている情報を利用することが可能となる [1]。本研究ではスライド作成支援のために、スライド間構造を考慮した再利用に適したスライド群の検索と発見の支援を行う。再利用に適したスライド群とは作成目的に合致するスライド間構造を持ったスライド群である。ユーザがスライド作成時にスライド間構造を考慮して検索できるスライドリポジトリを構築し、ユーザのスライド作成支援を行う。

3 スライドシーン

スライドシーンとは、意味を持ったスライド間構造の集合のことである。プレゼンテーション資料は様々な目的で作成される。研究発表や講義、メモなどが考えられる。研究発表が目的であるプレゼンテーション資料を考えたとき、スライド間構造が、“背景”、“目的”、“提案”などまとまりに分けることが可能となる。本研究ではこのスライド間構造をスライドシーンとし、プレゼンテーション資料からスライドシーンを抽出することを試みる。

プレゼンテーション資料 p はスライド $s_1 \dots s_n$ の n 枚で構成される。

$$p = s_1 \cdot s_2 \cdots s_{n-1} \cdot s_n (n \geq 1)$$

ある事柄について連続したスライドで述べているかたまりをチャンク c_i と呼ぶ。チャンクは共通する事柄について述べている連続したスライド群である。チャンク c がスライド s_q から連続するスライド k 枚で構成されるとき、次のように表される。

$$c = s_q \cdot s_{q+1} \cdots s_{q+k-2} \cdot s_{q+k-1}$$

またプレゼンテーション資料 p はチャンク $c_1 \cdot c_2 \cdots c_{m-1} \cdot c_m$ で表すことができる。

スライドシーンは再利用されやすいスライド間構造、すなわちチャンクの組み合わせである。本研究では繰り返し用いられるチャンクの組み合わせを特定し、スライドシーンとして利用する。またプレゼンテーション資料が作成された研究発表、講義、メモなどの目的のことをスライドシチュエーションと呼ぶ。スライド間構造である“背景”、“目的”、“提案”などの情報を各スライドにスライドシーンラベルとして付与する。

4 スライドリポジトリ

本研究ではプレゼンテーション資料をスライドシーンに基づいてスライドリポジトリに登録する。このスライドリポジトリを利用し、スライド作成支援を行うためのプレゼンテーション資料管理システムを試作する。スライドリポジトリにはスライドシーンの情報の他に、プレゼンテーション資料作成時間の時間情報を利用する。プレゼンテーション資料を入力すると自動でスライド間構造を判断し、スライドリポジトリに登録される。

4.1 スライドリポジトリの情報

プレゼンテーション資料 p は次のようにスライドリポジトリに登録される。プレゼンテーション資料 p はスライドの他に時間情報 t を保持している。

$$t = \text{time}(p)$$

プレゼンテーション資料の i 番目のスライド s_i にはスライドシーンラベル l が付与されている。スライドシーンラベルは予め定義しておく。

$$l = \text{label}(s_i), l \in L$$

またスライド s_i にはスライドシチュエーション A の情報とスライド本文 text も付加されている。よってスライドの属性 attribute は次のように表すことができる。

$$\text{attribute}(s_i) = (l, a, \text{text}), a \in A$$

スライドシチュエーション a におけるスライドシーンラベル l のスライドを持つスライド集合のチャンクを $c^{(l,a)}$ のように表す。このチャンクの組み合わせがスライドシーンとなる。

A Construction Method of Slide Repository Based on Slide Scene
Tatsuya DOI, Ken SHIMIZU, Taiki ITO, Hiroaki KAKIMOTO,
Shun SHIRAMATSU, Tadachika OZONO, and Toramatsu SHINTANI
Dept. of Computer Science and Engineering, Nagoya Institute of
Technology, Dept. of Computer Science and Engineering, Graduate
School of Ennggeering, Nagoya Institute of Technology.

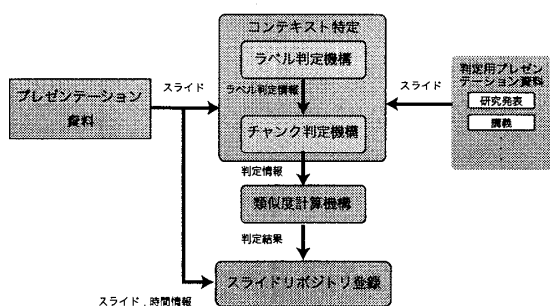


図 1: スライドリポジトリへの登録

4.2 スライドリポジトリへの登録

プレゼンテーション資料をスライド間構造に基づいて、スライドリポジトリへ登録を行う流れを図 1 に示す。

スライドシーン検出にはコンテキスト特定のための判定用プレゼンテーション資料を用いる。判定用プレゼンテーション資料はスライドシチュエーション毎に予め登録しておく。判定用プレゼンテーション資料には、スライドシチュエーションの判定基準となるスライド $s_{a,1} \dots s_{a,m}$ とスライドに付与されているスライドシーンラベル、スライド本文が登録されている。

まずラベル判定機構にスライド $s_1 \dots s_n$ を渡す。ラベル判定機構はスライドシチュエーション a の判定用プレゼンテーション資料 p_a のスライド $s_{a,1} \dots s_{a,m}$ と渡されたスライド s_j ($1 \leq j \leq n$) との類似度を計算する。スライド s_j と最も類似度の高かったスライド $s_{a,i} \in p_a$ のスライドシーンラベルをスライド s_j の (l,a) として判定する。各スライド s_j に対してスライドシチュエーション毎のプレゼンテーション資料 p と類似度計算を行い、スライドシーンラベルを判定する。判定結果はラベル判定情報としてチャンク判定機構に渡す。

チャンク判定機構はラベル判定情報をもとに、チャンク判定を行う。スライドシーンラベル $label(s_j)$ とスライド本文 $text(s_j)$ をから各スライドシチュエーション a のどのチャンクに属するのかを判定する。

チャンク判定においてすべてのチャンクに対してコンテキストの判定後、その判定情報を類似度計算機構に渡す。

類似度計算機構ではスライドシチュエーションをもとに、最終的なプレゼンテーション資料 p のスライドシチュエーション a_p を判定する。プレゼンテーション資料のスライドの属性が確定されるため、判定結果とスライドをスライドリポジトリに登録する。

4.3 プレゼンテーション資料検索

スライドリポジトリを利用したプレゼンテーション資料検索システムを試作した。本システムを利用することにより、PC 内に保存されているプレゼンテーション資料を収集し、再利用しやすい形に整理することが可能となる。スライド間構造を元にプレゼンテーション資料を分析しているため、スライドを検索する際にキーワードによる絞り込みとともにスライドシーンによるスライドの絞り込みを行うことが可能となり、より場面に適したスライドを提示される。本システムではスライドリポジトリから必要とするスライドを検索するための機能をもつ。スライドリポジトリを利用してスライドの検索を行った例を図 2 に示す。

図 2 は「web ページ」に関するスライドで、2008/10/01 から 2009/12/31 までに作成された「背景」、「目的」、「実験」のスライドシーン」の条件に該当する検索結果を示している。

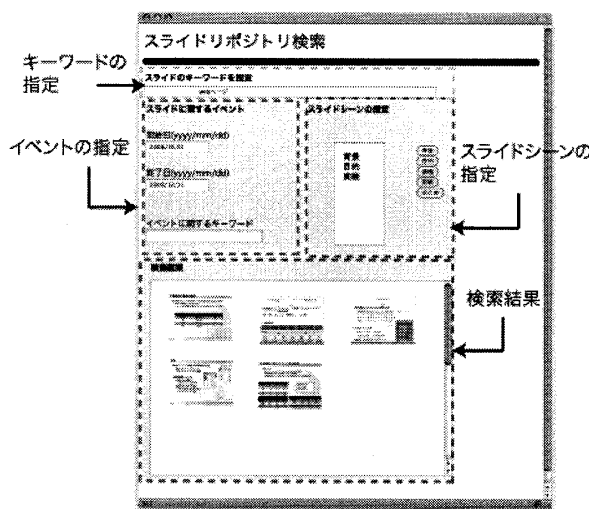


図 2: プレゼンテーション資料検索

スライドリポジトリ検索は 3 つの条件の指定が可能である。キーワード、イベント、スライドシーンの指定である。キーワードは検索結果に出力されそうなスライドの本文にありそうなキーワードを指定する。イベントはプレゼンテーション資料が作成された時期を指定するものである。イベントの指定は作成日時範囲の指定と、イベントに関するキーワードの 2 種類ある。イベントに関するキーワードは全国大会名など、イベントを特定できるものを入力する。スライドシーンは背景、目的、課題などのスライドシーンを選択する。右側のボタンから必要なスライドシーンを選択肢、左側のボックスに入力する。何も入力されていない場合は、すべてのスライドシーンが選択されたものとして検索される。

検索結果はスライドシーン単位となり、スライドのスクリーンショットが表示される。スライドシーンに含まれているスライドを閲覧したい場合はスクリーンショットをクリックするとポップアップウィンドウが表示され、スライドシーン内のスライドを確認することができる。

4.4 実装環境

本研究ではプレゼンテーション資料を PowerPoint ファイルとしている。対象となるファイルは拡張子が、「.ppt」、「.pptx」となるものである。プレゼンテーション資料管理システムは Mac OS 10.5 の PC で動作をするものである。

5 おわりに

プレゼンテーション資料の再利用支援としてスライドリポジトリ構築手法について提案した。スライドシーンに基づいてスライドリポジトリを構築することで、プレゼンテーション資料の流れを考慮したスライドの再利用法を利用することができる。本研究では研究発表と講義のプレゼンテーション資料に絞っているため、他のスライドシチュエーションを考慮したデータを取り入れる必要がある。

参考文献

- [1] 羽山 徹彩, 難波 英嗣, 國藤 進: “プレゼンテーションスライド情報の構造抽出”, 電子情報通信学会論文誌. D, 情報・システム. J92-D, pp. 1483-1494, 2009.
- [2] 清水 堅, 土井 達也, 佐野 博之, 工藤 聖広, 白松 俊, 大園 志親, 新谷 虎松: “スライドシーンに基づくスライド作成支援システムの実現”, 第 72 回情報処理学会全国大会, Mar, 2010 (掲載予定).