

学習支援のための検索意図に基づく教材スライドの抽出

山田歌織子[†] 王 元元^{††} 河合由起子^{†††} 荻野 哲男[†] 角谷 和俊[†]

[†] 関西学院大学総合政策学部 〒 669-1337 兵庫県三田市学園 2 丁目 1

^{††} 山口大学大学院創成科学研究科 〒 775-8611 山口県宇部市常盤台 2-16-1

^{†††} 京都産業大学コンピュータ理工学部 〒 603-8555 京都府京都市北区上賀茂本山

E-mail: †{fpm23477,togino77,sumiya}@kwansei.ac.jp, ††y.wang@yamaguchi-u.ac.jp,

†††kawai@cc.kyoto-su.ac.jp

あらまし 近年教育の現場でプレゼンテーション資料が教材として用いられる。しかしながら、スライドファイルがアップロードされる Web 上からユーザが必要とするスライドを取得することは困難である。そこで、本研究ではスライドを構造的・意味的に分析し、その結果からスライド検索のための演算を定義する。また、演算を用いた対話型の抽出手法を提案することで、ユーザの検索意図に合致するスライド抽出を可能とする。

キーワード e-ラーニング, 情報抽出, 教育

Kaoriko YAMADA[†], Yuanyuan WANG^{††}, Yukiko KAWAI^{†††}, Tetsuo OGINO[†], and Kazutoshi SUMIYA[†]

[†] Kwansei Gakuin University Gakuen, Sanda-shi, Hyogo 669-1337, Japan

^{††} Yamaguchi University Tokiwadai, Ube-shi, Yamaguchi 775-8611, Japan

^{†††} Kyoto Sangyo University Kamigamomotoyama, Kita-ku, Kyoto-shi, Kyoto 603-8555, Japan

E-mail: †{fpm23477,togino77,sumiya}@kwansei.ac.jp, ††y.wang@yamaguchi-u.ac.jp,

†††kawai@cc.kyoto-su.ac.jp

1. はじめに

近年教育の現場でプレゼンテーション資料(以下スライドファイル)が教材として用いられ、slide share に代表されるようなスライド閲覧サイト上にアップロードされる。スライドを閲覧する学習者は、学習内容に基づきスライドファイルもしくはスライドファイルを構成しているスライド 1 枚(以下スライド)を検索するために、スライド共有サイト内で検索を行うが、従来の検索機能においては、ユーザが入力したキーワードを AND・OR で組み合わせたクエリ入力による検索が用いられている。

しかし、このようなユーザによる入力には表記ゆれ等が含まれ、正確にスライドを検索することは困難である。また、ユーザは検索意図に合致するスライドが結果として出力されるまで、クエリの入力を繰り返さなければならない。そのため、スライドを推薦することで学習支援につながると考える。

2. 関連研究

講義資料として使われるスライドファイルなどを検索する研究は数多く存在する。片山ら [1] は、スライドコンテンツなど

に、講義を行った際の動作履歴を付与することで、学生が必要な部分だけを復習するための柔軟な検索を可能にしている。プレゼンテーションコンテンツの検索として、岡本ら [2] は、スライド構造やスライドの提示時間などのコンテキストを活用し、そのコンテンツ特性を考慮したシーン検索を可能にしている。また、友安ら [3] は、ポスターの閲覧には、閲覧者がポスター上のどの部分から閲覧すべきか即座にわからないという問題と、ポスターの構成要素同士の関係が明示的ではないという問題が存在すると分析している。そのためポスターとその構成要素の一部といえるスライドの構造分析に基づき、閲覧者のナビゲーションのためのズームング効果を用いたトランジションを自動的に生成するシステムを提案した。

王ら [4] は、インターネット上に公開されるプレゼンテーションコンテンツをユーザが閲覧する際、スライド間の概念的な詳細さを把握するため、キーワードの概念構造を利用したスライド間の関係判定方式について提案を行った。Wang ら [5] は、プレゼンテーションスライドの概要を把握するために、スライドを使用して iPoster と呼ばれるポスターを生成する方法を提案した。王ら [6] は、プレゼンテーションコンテンツを閲覧する



図 1 ユーザインターフェース

際のスライド間の関係に基づく要約や補足を実現するために、意味的構造と表層的構造を用いたスライド間の関係判定方式を提案した。

3. 提案手法

また、本研究では、ユーザとの対話型での検索手法を提案する。あるスライドをユーザが見ていた時に、そのスライドとユーザが過去に見たスライドの閲覧履歴を分析することで、そのユーザーの意図を読み取り、ユーザが次に閲覧すると推測されるスライドを複数個推薦する。また、スライドを抽出するための演算を定義する。演算を用いた対話型の抽出手法を提案することで、ユーザの検索意図に合致するスライド抽出を可能とする。

図 1 は本研究で提案するユーザインターフェースを示している。中央に閲覧中のスライドが表示され、その右側に、過去の閲覧履歴に基づき、次に閲覧されると推測されるスライドが推薦される。ユーザが選択した推薦ウィンドウ内のスライドが画面中央に表示される。

3.1 スライド間の関係

スライドを推薦と学習者の履歴を分析するにあたり、スライドの前後関係を検討する。王らの論文 [6] を用いてスライド間の関係を出しておく。

3.2 ユーザの検索意図

閲覧履歴からそのユーザが次にどんなスライドが見たいのかを検討する。本研究では「○×△のスライドを見たらつ次は■のスライドを推薦する」というルールを定める。

1. 詳細化

現在閲覧しているスライドファイルの文字数平均値より多い時、かつ履歴の直近 5 枚のうち 3 枚以上閲覧していた場合に詳細化のスライドを推薦する。

2. 簡略化

現在閲覧しているスライドファイルの文字数平均値より少ない時、かつ履歴の直近 5 枚のうち 3 枚以上閲覧していた場合に簡略化のスライドを推薦する。

3. 具体化

スライド間の関係が上位のスライドであり、履歴の直近 5 枚のうち 3 枚以上閲覧していた場合に具体化のスライドを推薦する。

4. 抽象化

スライド間の関係が下位のスライドであり、履歴の直近 5 枚のうち 3 枚以上閲覧していた場合に抽象化のスライドを推薦する。

3.3 検索演算

スライドを抽出するための演算の一覧を以下に示す。

- 「A」: タイトルが A のスライド
- [A]: 見出しが A のスライド
- A+B: A と B どちらかがでてくるスライド (OR)
- A-B: B が現れず A がでてくるスライド
- A × B: A と B 両方ででてくるスライド (AND)
- A → B: B よりも A が先にでてくるスライド
- A vs B: A と B を比較するスライド
- A > B: キーワード A が B より上部階層

この演算は、ユーザ自身が直接クエリとして入力することも可能だが、ユーザの過去に閲覧したスライドの履歴から、次に見るスライドとして推薦するためにスライドを抽出するためにシステム内部でも使用される。

3.4 推薦スライドの抽出

学習するユーザの閲覧履歴から、スライド間の関係と検索演算を組み合わせてスライドを抽出し、推薦する。

1. 詳細化

現在閲覧しているスライドの文字数 < and 類似度高類似度は TF-IDF を用いて類似度の高いものを算出する。

2. 簡略化

現在閲覧しているスライドの文字数 < and 類似度高類似度は TF-IDF を用いて類似度の高いものを算出する。

3. 具体化

先行研究よりスライド間の関係が下位のスライドを推薦する。

4. 抽象化

先行研究よりスライド間の関係が上位のスライドを推薦する。

4. おわりに

本研究では、対話型でのスライド抽出手法を提案した。提案手法として、スライドの記述方法とユーザの検索意図の分析、検索演算について定義した。

文 献

- [1] 片山薫, 香川修見, 神谷泰宏, 對馬英樹, 吉廣卓哉, 上林彌彦「遠隔教育のための柔軟な講義検索手法, 情報処理学会論文誌」Vol.39, No.10, pp2837-2854, 1998
- [2] 岡本拓明, 仲野亘, 小林隆志, 直井聡, 横田治夫, 岩野公司, 古井貞熙「プレゼンテーション蓄積検索システムにおける講義・講演音声情報を利用した適合度の改善」第 17 回データ工学ワークショップ (DEWS2006) 論文集, 6C-o1(2006)
- [3] 友安航太, 王元元, 角谷和俊「ポスターとスライドの構造に基づくズームングを用いたポスター閲覧方式」情報処理学会研究報告 Vol.2012-DBS-155 No.1
- [4] 王元元, 北山大輔, 角谷和俊「プレゼンテーションコンテンツのための概念構造に基づくスライドの関係判定方式」DEIM Forum 2010 C1-1
- [5] Yuanyuan Wang, Yukiko Kawai, Kazutoshi Sumiya「iPoster: Interactive Poster Generation based on Topic Structure and Slide Presentation」人工知能学会論文誌 30

巻1号 SPI-F(2015年)

- [6] 王元元, 北山大輔, 角谷和俊「プレゼンテーションコンテンツのための概念的構造と表層的構造に基づくスライドの関係判定方式」電子情報通信学会データ工学研究専門委員会, 日本データベース学会, 情報処理学会データベースシステム研究会 第2回データ工学と情報マネジメントに関するフォーラム (DEIM Forum 2010)