

文章の構造がレイアウトに反映されたスライドの自動生成

川西 哲平[†]
兵庫県立大学[†]

川嶋宏彰[‡]
兵庫県立大学[‡]

1 はじめに

スライドは簡潔かつ素早く情報を伝えられる手段として、会議や授業などの場面で広く使用されている。しかしスライド作成の作業は多くの時間と労力が必要となる。例えば、スライドのもととなるコンテンツの内容を十分理解し、掲載するテキストなどを考え、わかりやすくスライドにレイアウトすることが求められる。

これまでにドキュメントからスライドを自動生成する研究は行われているが [1, 2], これらの研究には2つの課題がある。1つ目は、スライドの主要な要素であるテキストが、どれも等しいレベルに並べられた箇条書きであるということ、2つ目は、スライドに含まれるテキストの色や大きさがすべて同じであるということである。実際に人が作成するスライドでは、箇条書きにインデントをつけたり、テキストボックスをレイアウトしたりすることが一般的である。また、作成者が伝えたいと思うフレーズや重要な単語は、色の変化や文字の太さ、大きさなどで強調されるべきである。

本研究ではこれらの課題を解決するために、キーワード抽出と談話関係分類を組み込んだスライド生成システムを提案する。キーワード抽出とは、文書中からその文書をよく表現している句（フレーズ）を抽出する手法である。また談話関係分類とは、2つの文の間にある関係を予測する手法であり、本研究においては接続詞予測を介した関係分類を行っている。これらの手法を組み合わせることで、既存研究の課題を解決したスライドの自動生成を目指す。

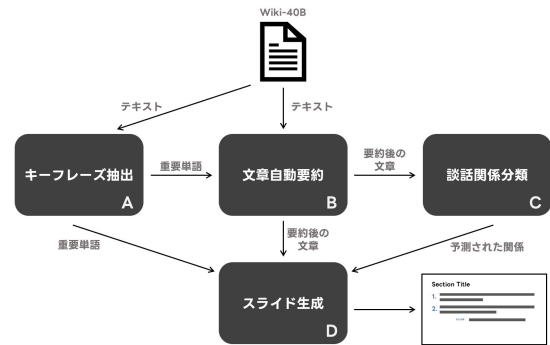


図1 提案手法の全体図

2 提案手法

提案手法は、図1に示すように大きく4つのモジュールから構成される。入力は英語 Wikipedia の1つの項目とし、データセット Wiki-40B [3] のテキストを用いる。出力は自動生成されたスライドであり、日本語訳まで行われるものとする。

最初に Wiki-40B のテキストをキーワード抽出モジュール（図1(A)）に入力する。次に、同じ Wiki-40B のテキスト、セクションタイトル、(A) のキーワード抽出モジュールで予測されたフレーズを文章自動要約モジュール (B) に入力する。得られた自動要約文章は、後段の談話関係分類モジュール (C) に入力し、文章中の2文間の関係を談話関係分類によって予測する。最後に、以上3つのモジュールで得た、キーワード、要約文章、文章関係を用いてスライドを生成する (D)。

2.1 キーフレーズ抽出

入力テキスト中の重要フレーズを抽出するために本研究では MultipartiteRank [4] を採用し、抽出結果の上位5フレーズを用いる。出力される5つのフレーズは、要約モデルの入力の一部、および出力スライド内でのフレーズの強調という2つの用途に用いる。

Sentence-Structure-Aware Automatic Slide Generation

[†] Teppei Kawanishi, University of Hyogo

[‡] Hiroaki Kawashima, University of Hyogo

2.2 文章自動要約

このモジュール (図 1(B)) では、既存手法である D2S [2] を用いて文章自動要約を行う。D2S とは、Question と Keyword で構成されるクエリを用いてクローズド・ドメインの長文形式の質問応答を行うことにより、文章からスライドに適した要約を生成するモデルである。本研究では、クエリの主要要素である Question にセクションのタイトル、それを補助する役割の Keyword に 2.1 節の出力結果を用いる。

2.3 談話関係分類

2.2 節の要約で得られる複数文の間の構造を、談話関係分類モデルにより推定する (図 1(C))。分類モデルには、Wikipedia より抽出されたデータセット [5] でファインチューニングした BERT を用いる。このデータセットには 20 個の接続詞が含まれるが、スライドのレイアウトに必要なのは接続詞そのものではなく、その接続詞から得られる文と文の関係である。そこで本研究では、談話関係分類などでよく用いられるデータセット The Penn Discourse Treebank 3.0 [6] のデータ構成を参考に、意味が重複した接続詞を同じ関係に落としこむなどして、分類クラスを 14 クラスに絞った。

2.4 スライド生成

スライド生成 (図 1(D)) は以下の手順で行う。まず、要約文章 (2.2 節) を関係分類の結果 (2.3 節) に基づき、あらかじめ設定したレイアウトを用いてスライド内に配置する。このとき同時に日本語訳を行う。その後、キーフレーズ抽出結果 (2.1 節) と文章内で一致する箇所の文字色を変更する。入力データである Wiki-40B の各セクションに対して一枚のスライドを生成し、スライド上部に表示されるタイトルはセクションタイトルとする。

3 結果

談話関係分類 使用したデータセット [5] のテストデータを用いて、2.3 節の方法で作成したモデルと、[5] の提案モデル (Decomposable Attention Model) との比較評価を行った。提案手法は 14 クラスでの分類を行うが、[5] と同条件にするために 20 クラスでの分類結果も併せ

表 1 談話関係分類精度の比較

Model	Macro F1	Accuracy
提案手法 (14 クラス)	0.586	0.589
提案手法 (20 クラス)	0.478	0.488
既存手法 (20 クラス) [5]	0.318	0.327

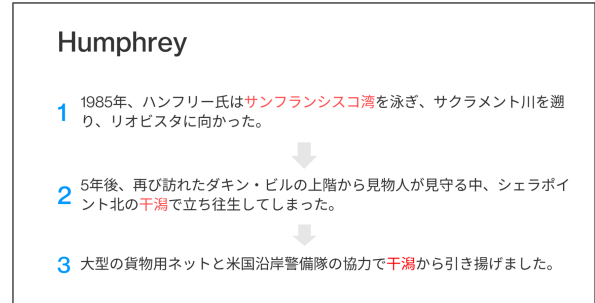


図 2 生成スライドの例

て表 1 に示す。本研究で作成したモデルは、いずれの指標においても [5] よりも精度が高く、14 クラスではさらに高い精度で分類可能である。

スライド生成 2.4 節の手法によるスライド生成結果の例を図 2 に示す。赤字強調のほか、このスライドでは「2 文目は 1 文目の後に起こる」、「3 文目は 2 文目の後に起こる」という各文の間に時間の流れを持つ関係が談話関係分類モデルにより推定され、それに沿って適切にレイアウト生成されていることがわかる。

謝辞 本研究の一部は科研費 JP19H04226 の補助を受けて行った。

参考文献

- [1] Tsu-Jui Fu, William Yang Wang, Daniel McDuff, and Yale Song. DOC2PPT: Automatic Presentation Slides Generation from Scientific Documents. *Proc. of the AAAI Conference on Artificial Intelligence*, 2022.
- [2] Edward Sun, Yufang Hou, Dakuo Wang, Yunfeng Zhang, and Nancy X. R. Wang. D2S: Document-to-Slide Generation Via Query-Based Text Summarization, May 2021. arXiv:2105.03664 [cs].
- [3] Mandy Guo, Zihang Dai, Denny Vrandečić, and Rami Al-Rfou. Wiki-40B: Multilingual Language Model Dataset. *Proc. of the 12th Language Resources and Evaluation Conference*, pp. 2440–2452, May 2020.
- [4] Florian Boudin. Unsupervised Keyphrase Extraction with Multipartite Graphs, April 2018. arXiv:1803.08721 [cs].
- [5] Eric Malmi, Daniele Pighin, Sebastian Krause, and Mikhail Kozhevnikov. Automatic Prediction of Discourse Connectives, Feb. 2018. arXiv:1702.00992 [cs].
- [6] Bonnie Webber, Rashmi Prasad, Alan Lee, and Aravind Joshi. The Penn Discourse Treebank 3.0 Annotation Manual. p. 81.