

議事録生成技術に関するサーベイ

Survey on Minutes Generating Technique

三浦 寛也^{1*} 平田 圭二¹
Hiroya Miura¹ Keiji Hirata¹

¹ 公立はこだて未来大学
¹ Future University Hakodate

Abstract: The purpose of this research is to consider the flow of the research and its problems by surveying the previous research on the minutes generation technology. Discuss how to proceed research on minutes generation techniques by clarifying the systematization of these conventional research and the position of each research. As a specific effort, we focused on the following four items and conducted a survey on minutes generation technology; summary generation, discourse structure analysis, structuring and visualization of discussion, and conversation analysis using multimodal corpus.

1 はじめに

本論文の目的は、議事録生成技術に関する従来研究をサーベイすることによって、当該研究における連綿たる技術の流れや、その問題点について検討することである。議事録とは、話し合った内容や会議での決定事項、出された意見等をまとめたものであり、会議の記録を残すためにはその作成が必須である。しかし全ての発言の中から重要な部分を見極め、それをまとめあげる作業には多大な労力と時間が要される。そのため、人間の手間をかけずに議事録を自動的に生成することが望まれている。近年の知識管理の分野においては、デジタル化されたマルチメディアデータを容易に取り扱うことが可能になり、音声・映像を議事録に組み込み可能なシステムも増えてきた [30, 36]。さらに、AMI Meeting Corpus [3] やディスカッションマイニングシステム [26] といった、会議録の効率的な閲覧に関する研究の支援を目的とした、人手により様々なアノテーション情報が付与された会議録データが提供されている。これらに含まれる情報量は膨大であるため、蓄積されたデータに対する多様な視点や、柔軟なものの見方への重要性が主張されており、この問題を解決するための技術が数多く提案されている。

本論文では、これらの従来研究の体系化および各研究の位置づけを明らかにすることで、議事録生成技術に関する研究の進め方について議論する。具体的な取り組みとして、以下の4項目に焦点を当て、議事録生成技術のサーベイを実施した；(1) 要約生成, (2) 談話



図 1: 本論文のアプローチ

構造解析, (3) 議論の構造化 (4) マルチモーダルコーパスを用いた会話分析 (図 1)。

2 要約生成に関する研究

会議録の主たる再利用法は議事録であり、議事録の内容を効率的に提示する要約は、会議録の閲覧にかかる時間を短縮し会議の流れや決定事項などの重要項目を効率よく把握できるようにすることが目的である。文書要約技術を応用したミーティングを対象とする要約生成の研究は、これまでに数多く行われている [7, 12, 22, 24]。自動要約のアプローチには、(1) 重要な発話を会議録中から取り出す抽出型の要約と、(2) 人が要約を行う際のように、会議録に含まれないような文を新たに生成する生成型の要約がある。一般に、抽出型要約では、テキスト解析による手法が広く用いられている [33]。重要文の抽出には、その文書のタイトルを利用する方法

*連絡先: 公立はこだて未来大学
〒041-8655 北海道函館市亀田中野町 116-2
E-mail: g3115002@fun.ac.jp

や、Support Vector Machine や LSI を用いて文の重要さの重み付けを行う方法など様々なアプローチが提案されている [8, 10, 16] . また、議事録自動生成の研究には、会議コーパスの会議会話履歴から議事録を生成する研究 [24] や、チャット会議ログから議事録を生成する研究 [14, 15] 等があるが、その多くは一般的な議事録のようにまとめた形ではなく、ほぼ抜き出した文そのままに近い形で重要文を提示している . これらの研究は、読むべき対象となる発話の量を大幅に削減することが可能であるが、発話構造や発話間の対応が考慮されておらず、構文的に不自然な文や意味を成さない要約が生成されるという問題がある .

複数文書要約研究では、利用者の要求や利用時の文脈によって重視する観点が異なるため、ある主題に関連した文書集合から内容を抽出して、利用者の要求に応じて重要な内容を提示することに主眼が置かれている . Tombros らは、利用者が検索要求として表現したクエリに焦点を当てて、要約を作成し分けることを提案した [34] . この手法では、トピックのように検索対象中に内容語として表現される要素に焦点を当てて要約を作成することができるが、事実、意見、知識などの情報のタイプという要素に焦点を当てて要約を作成し分けることはできない . この問題に対して、関らは、要約の観点として利用者が文書集合中のトピックと重視する情報のタイプとを指定する方法を提案した [25] . しかし、これらのアプローチで生成された要約は原文書が持つ談話構造を保持しているとは限らない . そのため、文間の依存関係を無視して要約が生成されることがあり得る .

3 談話構造解析に関する研究

議事録の生成においては、会議録における談話内の文間に存在する意味的関係を認識し、その構造を明らかにすることが必要である . 談話構造解析の分野においては、Marcu らの研究が代表的である [4, 18, 19] . この研究では、文書を修辞構造理論に基づく談話構造木として表現し、木構造に基づいて決定されるユニット間の半順序関係を利用した要約生成を実現している . 彼らは談話構造タグ付きのコーパスを作成し、機械学習の手法を用いて談話構造解析を行なった . これらの手法によって質の高い分析が実現されるが、談話構造タグ付きコーパスを作成するにはコストが掛かるといった問題が存在していた .

当該分野においては、Rhetorical Structure Theory (RST) [17] に基づく重要文抽出や [9, 21, 31] に関する研究が数多く行われている . 柴田らは、談話構造解析を用いて、入力テキストから要約スライドを自動生成する手法の提案し、その結果、入力テキストよりも見

やすいスライドを生成できることを確認した . この研究では、談話構造解析結果に基づき、抽出した主題部・非主題部を配置することによりスライドを生成している [32] . 会議録を対象とした談話構造解析に関する研究では、会議中の発話は自然発生的なものであるため、言い間違いなどによる不自然な単語の連続や構文的に正しくない言い回しなどが多く含まれている . そのため、会議録中の発話に対しては構文解析がうまく機能せず、これらの手法の応用は困難である .

4 議論の構造化に関する研究

会議録に残るテキスト情報を構造化することで、これらの情報を補完するという研究が存在する . Adbeen らは、テキストに対し構文解析や意味解析を適応することで、マインドマップを自動生成するシステムを提案した [1] . また、Elhoseiny らは、階層的な表示が可能なマインドマップの自動生成を行うシステムを提案した [6] . しかし、これらの研究はいずれも、構文的な解析が可能な書き言葉のテキストを対象としているため、会議録中の発話を対象として適用することは一般に難しい .

一方で、構文的な解析を用いない、会議録を対象とする効率的な閲覧システムに関する研究が存在する [20, 23, 35] . 松村らは、議事録に含まれる言語情報を用いて、トピックのセグメント分割を行い、セグメント単位による構造化されたマップの作成を行った [20] . 趙らは、議論の中で展開されるテーマは名詞の集合で表現できるという過程の元で、議事録内に含まれる名詞をノードとし、関係のあるノード間にエッジを貼った議論マップの自動生成を行った [5] . また、森らは、リフレクションのための発話間の関係に着目した議論構造モデルの提案および議事録の自動構造化手法の提案を行った [23] . これらの研究では、1つのトピックセグメントに含まれる文量が多すぎる点や、議論の流れに沿って正しく可視化しないと余計に見づらくなる点、ノードが名詞であるため分かりづらい点などの問題が挙げられている .

5 マルチモーダルコーパスを用いた会話分析に関する研究

先述した通り、自動要約や談話構造解析に関する従来手法では、一般にテキスト解析のみに注目していたため、発言場面の状況まで考慮できず、詳細な解析ができないという問題点があった . 一方で、社会学やコミュニケーション科学における会話分析の知見から、発話や行動の意味を正確に理解するには、対面会話中に

表 1: 議事録生成技術に関する研究の概要

分類	手法とその特徴	課題点
要約生成	<i>SVM</i> や <i>LSI</i> などの既存の言語処理技術を用いたテキスト解析	・発話構造や発話間の対応が考慮されていない
	会議ログに基づく重要文の抽出 トピックや情報の指定による異なる観点による要約の実現	・構文的に不自然な文や意味を成さないような文が多く生成される
談話解析	修辞構造理論に基づく談話構造木を用いた重要文抽出	・談話構造タグ付きコーパス作成のオーバーヘッド
	談話構造解析を用いた要約スライドの自動生成	・自然発生的な発話に対して構文的な解析がうまく機能しない
構造化	構文解析や意味解析によるマインドマップの自動生成	・書き言葉が対象である会議録中の発話への適用が不可能
	トピックのセグメント分割とセグメント間関係の構造化による議論構造の可視化	・議事録に含まれる全ての発話が確認すべき対象となる
	発話間の関係に着目した議論構造モデル	・正確な記述の必要性
会話分析	非言語特徴量を用いた会議状態の判別	・言語情報および非言語情報への考慮
	マルチモーダル情報を用いたグループ会話でのコミュニケーション能力の推定	・社会的なネットワークの文脈を考慮した分析への拡張
	議論参加者のコミュニケーション行動に基づく重要発言の推定	・コーパス作成の効率化

交わされる言語情報だけでなく、発話の前後の間合いや話者交替といった非言語情報の重要性を示唆している [2, 11, 13] .

このような背景から、人間の対話コミュニケーションと非言語情報との関係に関して様々な分野で研究が行われている。市野らは、会議状態を判別するために有効に作用する非言語情報の特徴量を明示した [11] . 岡田らは、グループ会話において参加者が表出する発話ターンや韻律といったマルチモーダル情報からコミュニケーション能力の推定を行った [28] . さらに二瓶らは、議論の自動要約のために含めるべき重要発言を議論参加者のコミュニケーション行動に基づき推定することを目的とした研究を行った [27] . これらの研究はすべて有効な特徴量を高い精度で明らかにしている。

ここで、Alex は、コミュニケーションにおけるアイデアの流れを把握するためには、ソーシャルネットワーク構造内の変数や、人々が互いに及ぼし合う社会的影響力の強さ、同意を示す社会的シグナルを考慮することの重要性を主張した [29] . そのため、これからの会話分析を対象とした研究では、言語・非言語情報を考慮するとともに、社会的なネットワークの文脈の中で個々の活動を理解することが必要であると考えられる。以上より、情報やアイデアの流れと人々の行動の間にある数理的関係性の定式化に取り組むことで、新しい議事録生成技術への貢献に寄与できると言えよう。

6 むすび

本論文では、議事録作成技術に関する研究として、従来研究の位置付けとその体系化を行った。先述した研究事例のまとめを表 1 に示す。筆者は、議事録作成技術のための要素技術やそれに基づき設計されるシステムは、概ねこれらのアプローチのいずれかに分類されると考えている。当該研究における問題の 1 つとして、評価方法が挙げられる。一般に、要約生成技術の多くは、要点を網羅することが重要であるとの考えのもと、再現率重視の評価方法を採用している [27] が、正解データの作成方法や、新たな構造化手法に対するモデルの妥当性評価に関しては、議論の余地があると考えている。本論文で紹介した各分野やその研究事例は、各研究者が各々の達成すべき研究目的において何を重視したのかが反映されるものであり、特定の目的に対して適切な手法が唯一に定まるものではないと考えている。当該分野に関する研究に従事する研究者にとって、本論文での議論が一助となれば幸いである。

謝辞

本研究の一部は科研費 (基盤研究 (B)26280089) の助成を受けたものである。

参考文献

- [1] Mohammad Abdeen, R. El-Sahan, A. Ismaeil, and M. C.E. Yagoub, Direct automatic generation of mind maps from text with m2 gen, In Proceedings of TIC-STH, the IEEE Toronto International Conference on Science and Technology for Humanity, pp.95-99 (2009).
- [2] 坊農真弓, 高梨克也, 多人数インタラクション研究の方法 -言語・非言語コミュニケーション研究のための分析単位とその概念-, 人工知能学会誌, Vol.22, No.6, pp.838-845 (2007).
- [3] Jean Carletta, Simone Ashby, and Sebastien Bourban, The AMI meeting corpus: A pre-announcement, in Proceedings of MLMI, pp.28-39 (2005).
- [4] Lynn Carlson, Daniel Marcu, and Mary Ellen Okurowski, Building a discourse-tagged corpus in the framework of rhetorical structure theory, In Proceedings of SIGDIAL, the Workshop on Discourse and Dialogue, Vol.16, pp.1-10 (2001).
- [5] 趙雲超, 松村真宏, 谷内田正彦, 音声認識された議事録からの議事マップ自動生成, 人工知能学会全国大会論文集, Vol. 6, pp. 221-221 (2006).
- [6] Mohamed Elhoseiny and Ahmed Elgammal, Text to multi-level mindmaps, International Journal of Multimedia Tools and Applications, Vol.75, No.8, pp.4217-4244 (2016).
- [7] Michel Galley, A skip-chain conditional random field for ranking meeting utterances by importance. In Proceedings of EMNLP, the Conference on Empirical Methods on Natural Language Processing, pages 364-372 (2006).
- [8] Yihong Gong and Xin Liu, Generic text summarization using relevance measure and latent semantic analysis, In Proceedings of SIGIR, the Annual International ACM Conference on Research and Development in Information Retrieval, pp.19-25 (2001).
- [9] Barbara J. Grosz and Candace L. Sidner, Attention, intentions, and the structure of discourse, Computational Linguistics, Vol.12, No.3, pp.175-204 (1986).
- [10] 平尾努, 磯崎秀樹, 前田英作, 松本裕治, Support Vector Machine を用いた重要文抽出法, 情報処理学会論文誌, Vol. 44, No.8, pp.2230-2243 (2003).
- [11] 市野順子, 田野俊一, 発言の時系列的パターンを用いた 会議における発散 / 収束の判別の可能性, 人工知能学会論文誌, Vol.25, No.3, pp.504-513 (2010).
- [12] Andreas Kathol and Gokhan Tur, Extracting question/answer pairs in multi-party meetings, In Proceedings of ICASSP, the IEEE International Conference on Acoustics, Speech and Signal Processing, pp.5053-5056 (2008).
- [13] Mark L. Knapp, Judith A. Hall, Terrence G. Horgan, Nonverbal Communication in Human Interaction, Cengage Learning (2013).
- [14] 小林竜己, チャット会議ログからの議事録自動生成 -領域分析と試作検討-, 言語・音声理解と対話処理研究会, SIG-SLUD-A202-02, pp.7-12 (2002).
- [15] 小林竜己, 談話の局所・中位構造を利用したチャット会議ログからの議事録自動生成, 言語・音声理解と対話処理研究会, SIG-SLUD-A203-05, pp.29-34 (2003).
- [16] H. P. Luhn, The automatic creation of literature abstracts, IBM Journal of Research and Development, Vol.2, No.2, pp.159-165 (1958).
- [17] William Mann and Sandra Thompson, Rhetorical structure theory: Towards a functional theory of text organization, Text Vol.8, No.3, pp.243-281(1988).
- [18] Daniel Marcu, A decision-based approach to rhetorical parsing, In Proceedings of ACL, the Annual Meeting of the Association for Computational Linguistics, pp.365-372(1999).
- [19] Daniel Marcu, The rhetorical parsing of unrestricted texts: A surface-based approach, Computational Linguistics , Vol.26, No.3, pp.395-448 (2000).
- [20] 松村真宏, 加藤優, 大澤幸生, 石塚満, 議論構造の可視化による論点の発見と理解, 日本ファジィ学会誌, Vol.15, No.5, pp.554-564 (2003).
- [21] Ryan McDonald, A study of global inference algorithms in multi-document summarization, In Proceedings of ECIR, the European Conference on Information Retrieval, pp.557-564(2007).
- [22] Yashar Mehdad, Giuseppe Carenini, Frank W. Tompa, and Raymond T. NG, Abstractive meeting summarization with entailment and fusion,

- In Proceedings of ENLG, the European Workshop on Natural Language Generation, pp.136-146 (2013).
- [23] 森幹彦, 八村太輔, 喜多一, リフレクションのための逐語議事録を用いた議論の構造化法, 人工知能学会全国大会論文集, No.2D4-1 (2007).
- [24] Gabriel Murray Abstractive meeting summarization as a markov decision process, *Advances in Artificial Intelligence* pp.212-219 (2015).
- [25] 関洋平, 江口浩二, 神門典子, 利用者の情報要求を考慮した観点に基づく複数文書要約とその評価, *情報処理学会論文誌*, Vol.46, No.8, pp.106-119 (2005).
- [26] Katashi Nagao, Katsuhiko Kaji, Daisuke Yamamoto, and Hironori Tomobe, Discussion mining: Annotation-based knowledge discovery from real world activities, In Proceedings of PCM, the Pacific Rim Conference on Multimedia, Springer Verlag Berlin Heidelberg, pp.522-531(2004).
- [27] 二瓶芙巳雄, 高瀬裕, 中野有紀子非言語情報に基づくグループ議論における重要発言の推定 グループ議論の要約生成に向けて , *電子情報通信学会論文誌 A* Vol.J100-A No.1 pp.34-44(2017).
- [28] 岡田将吾, 松儀良広, 中野有紀子, 林佑樹, 黄宏軒, 高瀬裕, 新田克己, マルチモーダル情報に基づくグループ会話におけるコミュニケーション能力の推定, *人工知能学会論文誌*, Vol31, No.6, AI30-E (2016).
- [29] Alex Pentland, *Social Physics: How Good Ideas Spread-The Lessons from a New Science*, Penguin Press (2014).
- [30] Lawrence A. Rowe and Vince Casalaina, Capturing conference presentations, *IEEE MultiMedia*, Vol.13, No.4, pp.76-84 (2006).
- [31] Athia Saelan Irfan Afif Filman Ferdian Ayu Purwarianti and Alfian Farizki Wicaksono, Natural language understanding tools with low language resource in building automatic indonesian mind map generator, *International Journal on Electrical Engineering and Informatics*, Vol.5, No.33, pp.256-269 (2013).
- [32] 柴田知秀, 黒橋禎夫, 談話構造解析に基づくスライドの自動生成, *自然言語処理*, Vol.13, No.3, pp.91-111 (2006).
- [33] Karen Sparck Jones, Automatic summarising: The state of the art, *Information Processing and Management: an International Journal*, Vol.43, No.6, pp.1449-1481 (2007).
- [34] Anastasios Tombros, Mark Sanderson, Advantages of Query Biased Summaries in Information Retrieval, In Proceedings SIGIR, the ACM Conference on Research and Development in Information Retrieval, Melbourne, Australia, pp.2-10 (1998).
- [35] 土田貴裕, 大平茂輝, 長尾確, 対面式会議コンテンツの作成と議論中におけるメタデータの可視化, *情報処理学会論文誌*, Vol.51, No.2, pp.404-416 (2010).
- [36] Alex Waibel, Tanja Schultz, Michael Bett, Robert Malkin, Ivica Rogina, Rainer Stiefelhaagen, and Jie Yang, SMaRT: The smart meeting room task at ISL, In Proceedings of ICASSP, the IEEE International Conference on Acoustics, Speech, and Signal Processing (2003).