

令和 5 年 6 月 7 日現在

機関番号：82626

研究種目：若手研究

研究期間：2020～2022

課題番号：20K19878

研究課題名（和文）作者の意図抽出と文章の自動生成を融合した包括的な作文支援システム

研究課題名（英文）Development of writer support system using the user's intention extraction and automatic text generation

研究代表者

渡邊 研斗（Watanabe, Kento）

国立研究開発法人産業技術総合研究所・情報・人間工学領域・研究員

研究者番号：50828324

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：歌詞に書きたい内容は決まっているがどんなフレーズで書けばよいか思いつかないユーザーに対して、ユーザーの意図を反映した歌詞を生成し提案することで、作詞の発想を与える支援システムを開発した。この支援システムの開発過程で、計算機が歌詞の意味内容を捉えるためには歌詞から字面に関する情報を取り除き、意味だけを抽出することが効果的であると判明した。そこで字面情報を取り除く方法を調査した結果、ユーザーが入力した文章を画像に変換することで、ユーザーの意図のみを抽出することが可能となり、その画像から歌詞を生成することで、意図を反映した歌詞生成の性能向上が確認された。更に著作物の剽窃リスクを下げる歌詞生成手法も開発した。

研究成果の学術的意義や社会的意義

本研究によって、画像を用いた情報抽出手法が歌詞生成の性能向上に有効であることも示され、自然言語処理の分野において、意味情報のみを取り出す手法の一つとして、有用な知見を提供することを意味する。また、本研究成果が音楽制作において、作詞者や音楽プロデューサーの創造性を支援することが期待される。特に、作詞に苦手意識を持つ人や初心者にとって、意図を反映した歌詞生成が手助けとなると考えられる。また、著作権侵害の問題が深刻化するGenerative AIにおいて、著作物の剽窃リスクを下げる歌詞生成手法が開発されたことは、AI業界と音楽業界にとっても大きな意義を持つ。

研究成果の概要（英文）：We have developed a lyric writing support system to assist users who have a clear idea of the content they want to include in their lyrics, but are struggling to come up with appropriate phrases. This system generates and suggests lyrics that reflect the user's intentions, providing inspiration for the creative process. In developing this support system, we found that removing information related to the wording of the lyrics and focusing only on the meaning was effective in enabling the computer to grasp the intended message. To accomplish this, we investigated methods for removing wording information and discovered that by converting the user's input text into an image, we could extract only the intended meaning. By generating text from this image, we were able to improve the accuracy of the user's intent in the generated text. In addition, we developed a lyric generation method that reduces the risk of plagiarism, which further enhances the usefulness of the system.

研究分野：自然言語処理

キーワード：歌詞情報処理 自然言語処理 音楽情報処理 ヒューマンコンピュータインタラクション 創作支援

1. 研究開始当初の背景

研究開始当初、深層学習による「文章の自動生成技術」や、90年代より盛んに議論された「発想支援技術」により、作文プロセスは簡略・高速化され、作文支援技術はさらなる発展を見せていた。特に自動生成技術を用いた作文支援では、作者が入力した未完成の作品や数値パラメータに対して尤もらしい単語列を探索することで、作文プロセスを効率化している。しかし、従来の自動生成による作文支援システムには以下の根本的な課題点がある。

- (1) 【計算機側の課題点】統計的機械学習による文の自動生成は、入力パラメータと出力文のパターンマッチを学習することにより「膨大な可能性の中から、条件に合致した尤もらしい単語列を探索する」ことはできるが、「生成された文の内容に一貫性・具体性を持たせる」ことは統計的な枠組みを利用するだけでは困難である。
- (2) 【人間側の課題点】「作品に盛り込むべき具体的かつ一貫した意図」を作者が適切なパラメータとして自動生成器に入力しない限り、(たとえ、文法や意味が完璧な文が自動生成可能だとしても)作者が望む文章は生成されないため、真に創作活動を支援できたとは言えない。

もし、作者の意図を計算可能なデータ構造として正確に抽出し、それを自動生成器に適切に自動入力できれば、計算機は「具体的かつ一貫した意図を条件とした文章の自動生成」が可能となり、人間は「自身の意図を正しく支援システムに伝える」ことが可能となる。つまり、本研究の問題設定は、作者から自身の意図を計算可能な形で抽出すること(Creator Intent Extraction)と、抽出された意図に基づいた文章生成(Intent-based Content Generation)、及びこれら技術の融合である。

2. 研究の目的

本研究の目的は、小説や詩、歌詞などの創作に関して、作者の意図を計算可能なデータ構造として抽出し、抽出された意図を生成文章に埋め込むことで、作品の表現力や作業効率がどこまで向上するのかを明らかにすることである。この目的達成のために、(1)作者の意図の表現形式の設計、(2)作者の意図を抽出するためのインタラクションの設計、(3)作者の意図を考慮した自動生成技術および作文支援システムの開発と評価に取り組む。

- (1) 作者の意図を表すデータ構造の設計：創作活動におけるアイデア整理を促進するために、従来の発想支援研究では、作品の概要やアイデアになりそうな情報を静的なグラフ構造として表現している。例えば、文章解析器を使った構文グラフ、Wikipedia等を用いた知識グラフ、単語ベクトルを用いた類似単語グラフ等がある。本研究では、これらの静的な情報に加え、「登場人物の感情」や「場面状況」のような動的に変化する情報を時系列関係構造として表現することで、作者の意図に具体性と一貫性を持たせる。また、これら静的構造と動的構造を大量のテキストデータから抽出する機械学習手法を考案し、各構造が作者の意図をどの程度まで表現できるかを明らかにする。
- (2) 作者の意図を抽出するためのインタラクションデザイン：従来の作文支援では、キーワードや、数値パラメータ、メモ書きなどの特定の情報を入力する「クエリ型インタラクション」によって、作者の意図を局所的に抽出している。本研究では、システムが出題する様々な質問に対して作者が回答を繰り返すことで、作者の意図構造を包括的に抽出し、その結果をフィードバックする「相談型インタラクション」を考案する。具体的には、作者にYes・No質問や記述質問などを答えてもらい、その回答内容から推定した意図グラフ構造を提示し、必要ならばグラフ構造を修正し、これらのプロセスを繰り返すことで、作者の意図を具体的に表現したデータ構造を効率的かつ正確に抽出する。
- (3) 作者の意図を考慮した自動文章生成・支援システムの有用性評価：作者の意図抽出技術と文章の自動生成技術を統合し、全ての作文プロセスを包括した総合的な作文支援インタフェースをWebサービスとして構築する。また、各機能が作者の作業効率と作品の表現力の向上にどれだけ寄与するかを明らかにする。

3. 研究の方法

当初の計画は次に示すようなものであった。まず、初年度に作者の意図を表現したグラフ構造を抽出する機械学習モデルを構築する前準備として、「作品内容を表したグラフ構造」と「グラフ構造を説明した質問文」のペアデータを作成する。ベイズ学習を用いたストーリー構造を推定する手法を用いて時系列関係構造を解析することから検討し、解析した時系列関係構造が詳細に読解可能であれば、ルールベース手法でも高品質な質問文を作成できると考えている。作成したデータの質を評価するために「グラフ構造に矛盾はないか」「グラフ構造を読解できるか」等の主観評価スキームを設計し、単純な構文グラフや類似単語グラフと比較することで、提案グラフ構造の表現能力を明らかにする。

二年目では、システムと作者が質問と回答を繰り返すことで作者の意図を具体的に包括的に抽出する「相談型インタラクション」を設計する。相談型インタラクションでは「どの質問をどの順番で作者に提示するか」「回答から目的の意図構造をどのように予測するか」の二つの技

術的課題を解決する必要がある。本研究では、作者の回答履歴から「最大の情報量を得られる次の質問を選定する」ことに加え「グラフ構造を徐々に推定する」ことを同時に行う数理モデルを構築する。モデル学習の際には、敵対的学習などの知見を参考にしながら、誤った質問文や回答が入力された場合に、正解とは異なるグラフ構造を識別するような目的関数の設計を行う。手法の有用性は、正解グラフの推定に至るまでの質問数や、正解グラフとの一致率を計測することで明らかにする。

そして最終年度に、「作者の意図を抽出するモジュール」と「意図に基づいた文書を生成するモジュール」を統合した作文支援インタフェースを実装する。本インタフェースは二つのモジュールを単に連結したのではなく、創作途中であってもユーザが自在に両モジュールを行き来できる柔軟性の高いものである。本インタフェースを Web サービスとして公開し、作業時間や作品内容を評価することで、創作過程がどこまで効率化・高速化され、作品の表現力がどこまで高められたかを明らかにする。

4. 研究成果

歌詞データは曖昧な表現が多く、フォーマットが統一されていないため、作者の意図を解析することが困難であることがわかった。そこで、2020 年度に予定されていた作者の意図を表現した時系列関係構造データの作成ではなく、歌詞生成技術を開発した。本技術は、創作的文章の創作支援を目的としており、作者の意図を反映した歌詞を生成することができる。従来の歌詞生成技術には、2つの課題があった。1つは、バラード楽曲とメタル楽曲などの曲調に合わせた歌詞生成ができなかったこと、もう1つは、ありがちな表現とありがちなでない表現の使い分けを考慮した歌詞生成ができなかった。そこで、曲調を表現した音響信号とありがちなでない単語を多次元ベクトルで表現し、これらが近くに配置されるようなベクトル空間を構築することで、音楽の雰囲気と正しい文法の両方を満たす単語を生成する作詞支援技術を開発した。(図1)

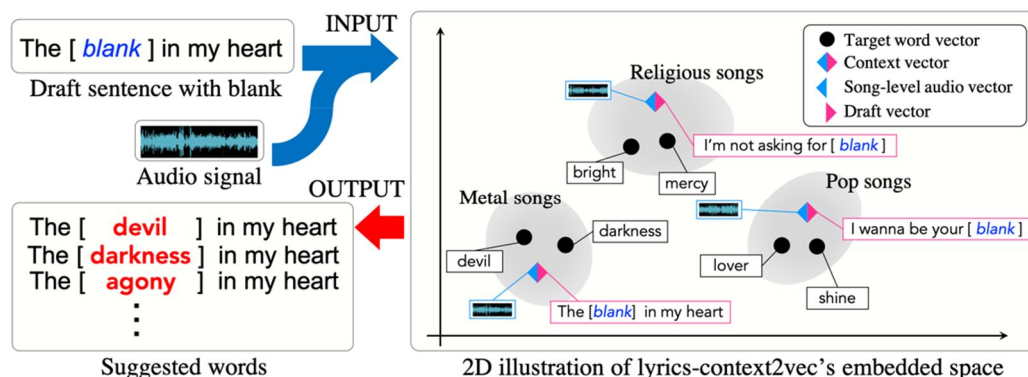


図 1：ありがちな歌詞生成の概要

実験の結果、従来の歌詞生成手法と比較して音響信号に対してありがちなでない単語の補完が可能になったことがわかった。さらに、音響信号と単語には相関関係があるということ、巨大な音楽データセットから定量的に観測できた点も学術的に価値のある結果となった。本研究の成果は国際会議 Multimedia Modeling に採択され[1]、また、言語処理と音楽情報処理の国際ワークショップである NLP4MuSA にて発表した。[2]

次年度には、歌詞を包括的に分析するための前処理技術開発に取り組んだ。歌詞は作家性が強いので、文章として不自然な箇所が改行/段落分けされることが多い。そのため最新の言語処理技術であっても精緻な分析が難しい。そこで自動段落分け、サビ箇所推定、文修正技術を開発した。特に、楽曲特有の繰り返し構造の特徴量に加え、強力な言語モデルである BERT を用いて言語的特徴量を計算することで、より高い精度でサビ箇所の推定可能とした。本成果は IF 付き国際論文誌 IEICE Transactions on Information and Systems に採択された。[3]

更に前年度で得た知見に基づき、歌詞を包括的に分析するために「歌詞の意味的内容の類似性」と「歌詞のスタイル(作者の作風や言葉遣い、文法的構造)の類似性」を捉える特徴量学習の開発に取り組んだ。具体的には、自然言語処理分野/音楽情報処理分野の両面から「歌詞および楽曲の類似度」に着目し、一千万曲以上の歌詞や音源データを用いて様々な機械学習モデルの構築及びパラメータの調整を行った。その膨大な調査の結果、従来の特徴量学習手法では歌詞の大まかな意味の類似性しか捉えられず、細かな意味の違いを捉えることは非常に困難であることがわかった。そこで、細かな意味の違いを捉えた歌詞特徴量を学習するために、教師データとして「意味は類似しているが、表層が異なる文ペア」を用意し、これらの文ペアを高次元ベクトル空間上の近い位置に配置するように学習した。この文ペアを用意するために、本研究では機械翻訳を用いて日本語 英語 日本語と逆翻訳することで、大量の教師データを自動獲得した。この教師データを用いることで、従来手法よりも細かな意味の違いを捉えた特徴量が学習可能となった。

最終年度では、前年度の「意味は類似しているが、表層が異なる文ペア」の分析結果から着想を得て、ユーザが入力した文章の意味は類似しているが、表層が異なる新たな歌詞を生成することで、ユーザに作詞の発想を与える支援システムの開発をした。具体的には、入力文章の字面に関する情報を取り除きつつ入力の意味内容のみを抽出するために、既存の画像生成技術を用い、その画像から歌詞を生成する機械学習手法を提案した。これにより、歌詞に書きたい内容はわかっているが、どうやって書けばよいかわからないユーザに対して歌詞を生成し、提示することで作詞の新たな発想を与えることが可能となる。(図2)



入力文章	中間表現(画像)	生成歌詞(4行制約付き)
二人の人が戦っていて、観客が観ている		このままじゃ 遊びは終わらせない 勝ち上がるだけさ! 全てを賭けて
ご機嫌なナンバーだぜ ラジオから流れてる 俺を急がせるよ Crazy DJ		このビートが刻むリズムの鼓動 これは音に乗っかって 俺らの鼓膜を揺らすんだろ? dj mc と mic で作れるかな?

図 2：入力文章に対して生成された画像および歌詞の例

更に、自動歌詞生成技術を用いた創作活動を円滑にするために、既存の歌詞を剽窃するリスクを下げた歌詞生成手法を開発した。本手法は「生成のために利用してもよいありがちなフレーズ」と「剽窃リスクがある特徴的なフレーズ」を統計的な観点から分類することで、学習データ中の特徴的なフレーズを原形的に生成す可能にした。本研究の成果は言語処理学会年次大会にて発表し、発表論文 579 件中 26 件(4%)に与えられる委員特別賞を受賞した。[4]

[1] Kento Watanabe and Masataka Goto. 2021, “Atypical Lyrics Completion Considering Musical Audio Signals”. In Proceedings of the 27th International Conference on Multimedia Modeling (MMM 2021), pp.174-186, (2021).

[2] Kento Watanabe and Masataka Goto. 2021. “Atypical Lyrics Completion Considering Musical Audio Signals”. In Proceedings of the 2nd Workshop on NLP on Music and Spoken Audio (NLP4MuSA 2021), pp.37-41, (2021).

[3] Kento Watanabe and Masataka Goto. 2023. “A Method To Detect Chorus Section in Lyrics text”. IEICE Transactions on Information and Systems, (2023) (To be appeared).

[4] 渡邊研斗, 後藤真孝. 2023. “入力文章の内容に沿った新たな歌詞を生成する作詞支援システムと剽窃リスクを下げる歌詞生成手法”, 言語処理学会第 29 回年次大会, (2023). 委員特別賞

5. 主な発表論文等

〔雑誌論文〕 計4件（うち査読付論文 4件 / うち国際共著 0件 / うちオープンアクセス 2件）

1. 著者名 Kento Watanabe and Masataka Goto	4. 巻 Vol. E106-D, No. 9
2. 論文標題 A Method To Detect Chorus Sections in Lyrics Text	5. 発行年 2023年
3. 雑誌名 IEICE Transactions on Information and Systems	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Kento Watanabe and Masataka Goto	4. 巻 1
2. 論文標題 Atypical Lyrics Completion Considering Musical Audio Signals	5. 発行年 2021年
3. 雑誌名 Proceedings of the 2nd Workshop on NLP for Music and Spoken Audio	6. 最初と最後の頁 37-41
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Kento Watanabe and Masataka Goto	4. 巻 1
2. 論文標題 Atypical Lyrics Completion Considering Musical Audio Signals	5. 発行年 2021年
3. 雑誌名 Proceedings of the 27th International Conference on Multimedia Modeling	6. 最初と最後の頁 174 ~ 186
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-67832-6_15	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Kento Watanabe and Masataka Goto	4. 巻 1
2. 論文標題 Lyrics Information Processing: Analysis, Generation, and Applications	5. 発行年 2020年
3. 雑誌名 Proceedings of the 1st Workshop on NLP for Music and Audio	6. 最初と最後の頁 6 ~ 11
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

〔学会発表〕 計4件（うち招待講演 0件 / うち国際学会 3件）

1. 発表者名 渡邊 研斗, 後藤 真孝
2. 発表標題 入力文章の内容に沿った新たな歌詞を生成する作詞支援システムと剽窃リスクを下げる歌詞生成手法
3. 学会等名 言語処理学会第29回年次大会
4. 発表年 2023年

1. 発表者名 Kento Watanabe and Masataka Goto
2. 発表標題 Atypical Lyrics Completion Considering Musical Audio Signals
3. 学会等名 the 2nd Workshop on NLP for Music and Spoken Audio (国際学会)
4. 発表年 2021年

1. 発表者名 Kento Watanabe and Masataka Goto
2. 発表標題 Atypical Lyrics Completion Considering Musical Audio Signals
3. 学会等名 The 27th International Conference on Multimedia Modeling (国際学会)
4. 発表年 2021年

1. 発表者名 Kento Watanabe and Masataka Goto
2. 発表標題 Lyrics Information Processing: Analysis, Generation, and Applications
3. 学会等名 The 1st Workshop on NLP for Music and Audio (国際学会)
4. 発表年 2020年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------