

# LYRICS RADAR: 歌詞の潜在的意味分析に基づく 歌詞検索インタフェース

佐々木 将人<sup>1,a)</sup> 吉井 和佳<sup>2</sup> 中野 倫靖<sup>2</sup> 後藤 真孝<sup>2</sup> 森島 繁生<sup>1</sup>

**概要:** 本稿では、歌詞検索インタフェース LYRICS RADAR について述べる。従来の歌詞中の語句に対する全文検索システムでは、ある単語（例：「涙」）をクエリとして入力すると、全く異なる意味の歌詞を持つ楽曲（例：失恋の「涙」と感動の「涙」）が混在した検索結果となり、ユーザの検索意図を十分反映できない問題があった。歌詞の意味をクエリとしての的確に言葉で表現して入力するのは困難なため、本研究では、クエリとして既存の歌詞を活用する検索インタフェースを提案する。具体的には、潜在的ディリクレ配分法を用いて多数の楽曲の歌詞から各楽曲の歌詞のもつ意味（トピックの比率）を推定することで、クエリとなる楽曲の歌詞と意味的に類似した歌詞をもつ楽曲が検索できる。歌詞のトピックの比率を五角形内に着色して可視化するトピックレーダーチャートや、データベース中のすべての歌詞をトピックの類似度に応じて二次元平面上にマッピングすることで、ユーザがトピックを言語表現する必要がなくなり、どのようなトピックの歌詞を探したいかわからない曖昧な状況でも活用できる。さらに、メタデータによる絞込などの機能によりユーザの求める未知の歌詞へ直感的に辿り着く方法を実現する。

## 1. はじめに

歌唱を伴うポピュラー音楽では歌詞が重要であり [1]、様々な気持ちや状況、情景が歌詞の中で表現されている。しかし、自分好みの歌詞を持つ楽曲を見つけることは容易でなく、楽曲を聴いてみないとその歌詞を気に入るかどうかわからないことが多い。歌詞中に登場する語句を表層的にテキスト全文検索することならば技術的に可能であるが、そこに表現されている気持ち等を検索できるわけではないので、有用性に限界がある。そこで本研究では、歌詞が潜在的に持つ意味（トピック）を自動的に分析することで、歌詞に基づく未知の楽曲との出会いを支援する歌詞検索インタフェースを実現することを目的とする。

従来、歌詞に基づいた楽曲検索や分類の手法として、歌詞のテキストを自然言語処理に基づいた手法で解析することで感情やジャンルなどで分類、マッピングを行う研究 [2][3][4][5] が挙げられる。また、歌詞の潜在的な意味に踏み込んで、楽曲の楽譜とその歌詞との関係をモデル化して、音符列と歌詞に基づく楽曲検索を可能にする研究 [6] や、楽曲の音響特徴空間と歌詞や関連 Web ページの語句

の特徴空間を対応付けることで、入力 Web ページに合った楽曲を提示する研究 [7] 等もあった。既存の歌詞検索サイトには、「恋愛」「卒業」のような人手で付与したタグに基づく検索機能もあるが、人手でのタグ付けには限界があり、同じタグが付与された楽曲が多ければ絞り込みが難しくかった。検索以外に歌詞を活用した音楽インタフェースとしては、歌詞と楽曲とを時間的に対応付けてカラオケ表示する LyricSynchronizer [8] や、歌唱の録音において歌詞を活用する VocaRefiner [9]、再生時刻指定や区間分割に歌詞を活用する合唱生成システム [10] 等があった。しかし、以上の従来の研究では、歌詞に関連した検索等はできても、ユーザがインタラクティブに自分好みの歌詞に自在にたどり着けるような技術は実現されていなかった。

そこで本研究では、歌詞が潜在的に持つ代表的な意味（トピック）を機械学習手法（潜在的ディリクレ配分法）で求め、多数の既存の歌詞の中から、ユーザが好む歌詞をインタラクティブに検索できる歌詞検索インタフェース「LYRICS RADAR」を提案する。同じ単語でも、使われ方によって複数の「トピック」がその裏に潜んでいる。例えば、メインボーカルが合成歌声の楽曲 4597 曲の歌詞に対してトピックを自動分析したところ、同じ単語「舞う」が、「踊る、回る、ドレス、ダンス」等の代表語をもつトピックと「桜、花びら、吹雪、葉」等の代表語をもつトピックの二つに属していた。代表語から推測すると、前者は「人が踊

<sup>1</sup> 早稲田大学  
Waseda University

<sup>2</sup> 産業技術総合研究所  
National Institute of Advanced Industrial Science and Technology (AIST)

<sup>a)</sup> joudanjanai-ss @ akane.waseda.jp

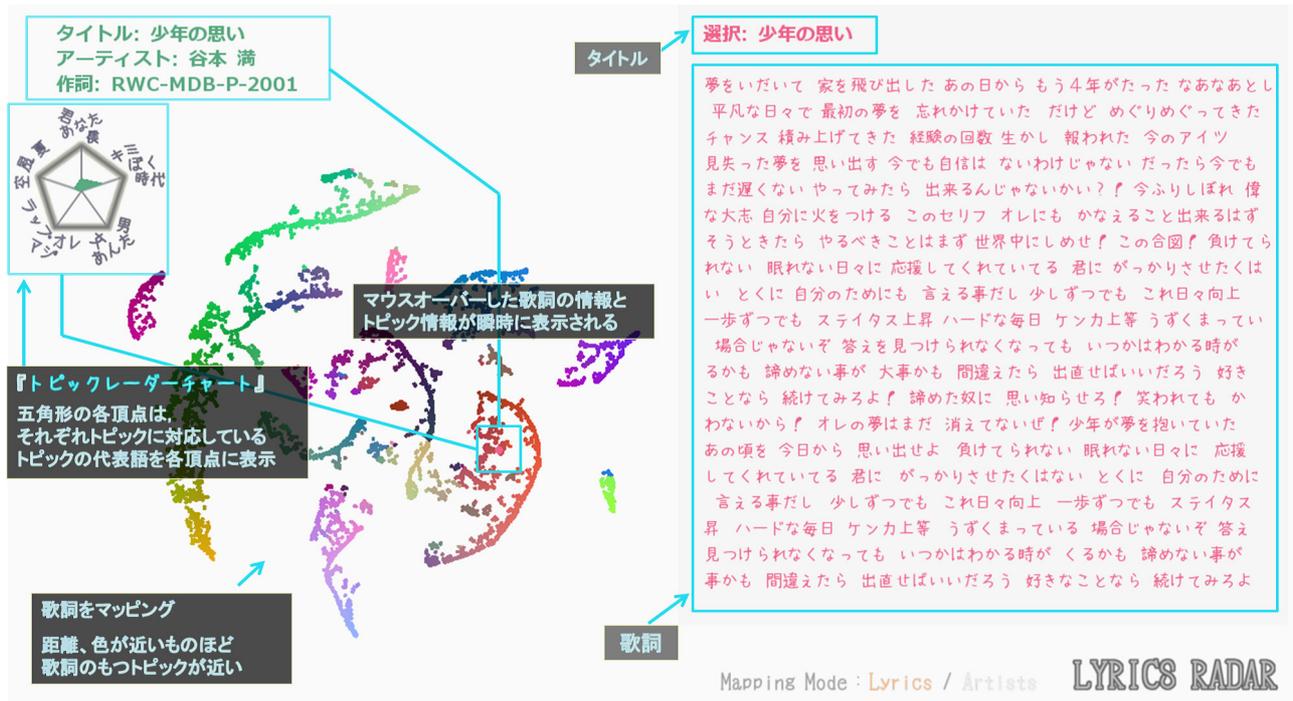


図 1 LYRICS RADAR の歌詞検索用インターフェースの表示例 (実際にはポピュラー音楽 (J-POP) を用いて実装しているが、本図では歌詞の例示のために RWC 研究用音楽データベースの楽曲 (RWC-MDB-P-2001 No.30) を用いた)

る動作」というトピックにおける「舞う」、後者は「花びらや雪などの軽い物体の動き」というトピックにおける「舞う」として使われていると考えられる。ただし、本研究ではこうした個別のトピックではなく、データベース中のすべての歌詞に共通する代表的なトピック 5 種類を求め、各歌詞をそれらの比率で表現する。この 5 個の値の比率を五角形の形状で表現した表示を「トピックレーダーチャート」と呼ぶ。ユーザがそれを見れば各歌詞に各トピックがどの程度関係しているかがわかり、直感的な歌詞検索・可視化が可能になる。

また、歌詞のトピックを推定するために、歌詞の選別及びアーティスト単位での推定を行った。歌詞は、楽曲で歌われる文字であるため、一楽曲の中で用いることの出来る単語の数は限られている。さらに、A メロやサビなど繰り返し使用される単語が存在するため、一般的な文章と比べると登場する単語数は少ない。トピック推定は、文章内の単語に依存するため、単に歌詞を文章とみなしてもトピックを正確に推定することは困難である。そこで、単語数の閾値の設定による歌詞の選別や、アーティスト単位でのトピック推定により歌詞の持つトピックの比率を求めた。

LYRICS RADAR では、トピックの比率が近い歌詞 (トピックレーダーチャートが似ている歌詞) が近くなるように、すべての歌詞を二次元平面上に配置する。ユーザは、この平面を覗くことで、好みの歌詞の近傍に位置するトピック比率の似た歌詞を発見できる。またユーザはトピックレーダーチャートの形状を直接変形させることで、歌詞

を検索することもできる。

## 2. 歌詞の潜在的な意味をクエリとして検索を行うインターフェース LYRICS RADAR

LYRICS RADAR は、潜在的ディリクレ配分法によって多数の楽曲の歌詞に共通して出現する複数のトピックを自動的に求め、各楽曲の歌詞のもつ潜在的な意味をトピックの比率で可視化したり、それをクエリとして活用したりすることで、直感的な検索を可能にするインターフェースである。これにより、単に歌詞中に出現する語句をクエリとした検索では実現できない、歌詞の意味に踏み込んだ新たなインタラクティブな歌詞検索を実現する。ユーザがトピックを言語表現する必要がないので、どのようなトピックの歌詞を探したいのかわからない曖昧な状況でも活用できる特長を持つ。

### 2.1 歌詞のトピックに基づく可視化機能

LYRICS RADAR は、トピックレーダーチャートと二次平面へのマッピングという 2 種類の可視化機能を持つ。図 1 にそのインターフェースの画面表示例を示す。トピックレーダーチャートは、各歌詞の潜在的な 5 種類のトピックの比率を五角形内に着色して可視化する機能であり、図 1 の左上に表示されている。マッピングは、データベース中のすべての歌詞をそのトピックの類似度に応じて二次元平面上にマッピングし、ユーザがインタラクティブに歌詞を探ることができる機能であり、図 1 の左側に表示されてい

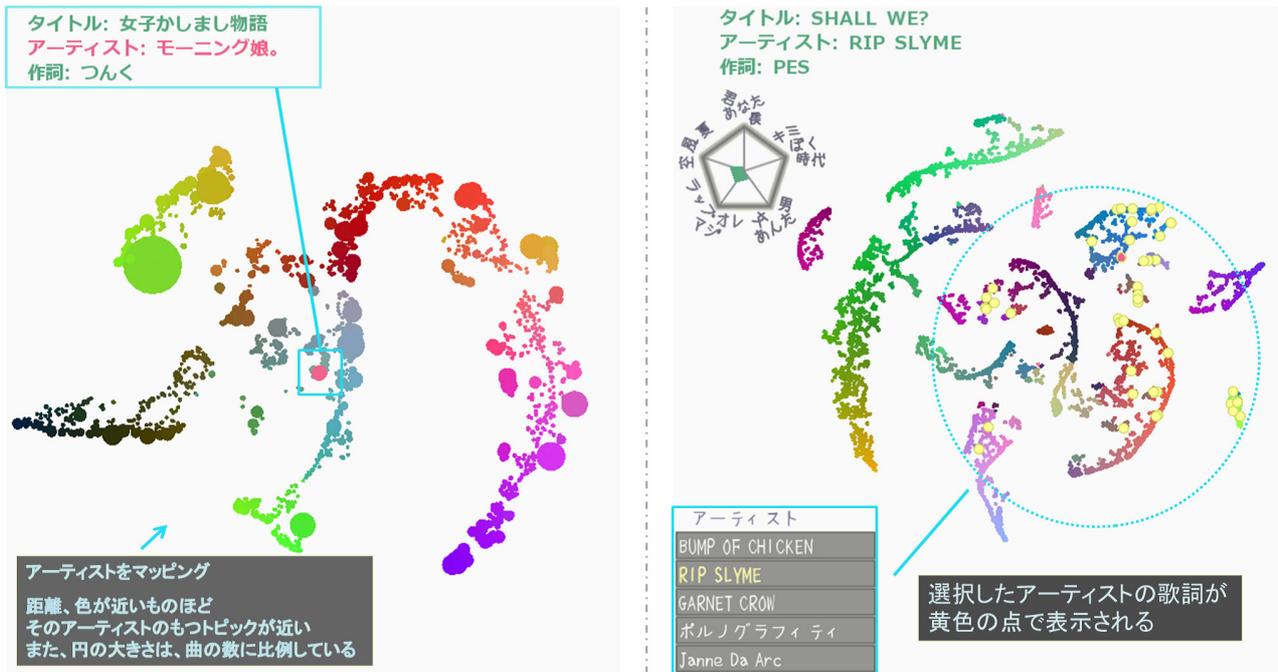


図2 アーティストのマッピング (左図), アーティスト名選択による歌詞の表示例 (右図)

る。また、アーティスト単位でトピック分析を行うことで、アーティストのマッピングも行った。トピックレーダーチャートが絶対的な指標に基づく可視化機能なのに対し、二次平面へのマッピングは相対的な指標に基づく可視化機能である。以下、それぞれについて説明する。

### 2.1.1 トピックレーダーチャート: 5種類のトピックの比率を表現した五角形

すべての歌詞に共通する5種類のトピックを求め、その各トピックの要素を各歌詞がどれぐらいの比率で含んでいるかを自動分析した結果を、五角形内にプロットして表示した可視化がトピックレーダーチャートである。五角形の各頂点が異なるトピックに対応し、図1の左上のように、そのトピックと関連の高い代表語を五角形の外側に複数表示することで、ユーザが各トピックの意味を類推できるようにした。

外周に近いほど大きな比率を表し、5個の値の合計が一定値であるため、ある一つのトピックの比率が突出していると、尖った形状になる。これにより、選択した歌詞のトピックを視覚的に把握しやすく、歌詞間の直感的な比較もしやすい。

なお、本インターフェースでのトピック数は、潜在的ディリクレ配分法で求めたトピックの内容と、インターフェースとしての操作性のバランスから5個に決めた。トピック数を増やせばより細分化した意味内容を扱うことができる可能性があるが、ユーザにとって操作がより煩雑になるトレードオフの関係にある。

### 2.1.2 歌詞をマッピングした二次元平面

トピック比率に近いほど配置が近くなるように各楽曲の

歌詞を二次元平面上にマッピングし、その一部を拡大して歌詞を探すことが出来る。各歌詞に対応する点は、各歌詞の持つトピック比率を三次元に圧縮し、それぞれRGBに対応させることで着色されている。これにより、歌詞の色を見ることで、トピックに基づいて歌詞がどのように分布しているかを一目で確認できる。マッピングや色付けにおける次元圧縮は、Maatenらが提案したt-SNEを用いた[11]。ユーザがカーソルをマウスオーバーすると、桃色に着色され、図1左上にタイトル、アーティスト名、作詞家名のメタ情報、その下にトピックレーダーチャート、右側に歌詞が表示される。これらは、カーソル移動でマウスオーバーを繰り返すことで、次々とリアルタイムに更新表示される。こうして、ある歌詞にトピックが類似した他の歌詞を発見することができる。マッピングされた歌詞は、ドラッグやキーボード操作で移動、拡大、縮小ができる。さらに、各歌詞にメタデータとして付与されているアーティスト名と作詞家名を活用した可視化も可能である。図2の右側のようにアーティスト名を選択すると、そのアーティストの歌詞の点が黄色で着色され、作詞家名を選択すると、その歌詞の点がオレンジ色で着色される。これは、アーティストや作詞家をクエリとした歌詞検索に相当するが、アーティスト、作詞家ごとにいかに分布しているかを直感的に把握できる点が新しい。従来の音楽情報検索では、アーティスト名での検索は活用されているが、作詞家名での検索が活用される機会は乏しかったが、本機能によって、自分の好きな歌詞の作詞家が手がけた他の歌詞に興味を持つきっかけが増え、新しい歌詞との出会いが広がる可能性がある。

また、アーティスト単位でトピック分析をすることで、

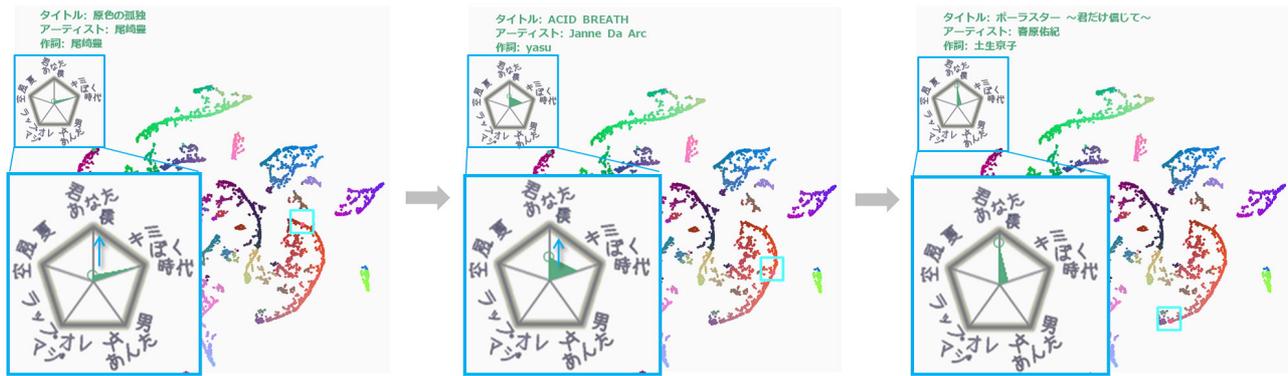


図 3 トピックレーダーチャート上でトピック比率を直接入力する具体例。  
五角形の上部の比率をマウスのドラッグにより上へ伸ばした様子を示す。

トピック比率が近いほど配置が近くなるようにアーティストを配置した。歌詞の場合と同様に、トピック比率に基づいて色付けされる。図 2 の左側のように、アーティストの持つ楽曲数に応じて、円の半径が大きく表示される。こうして、あるアーティストにトピックが類似した他のアーティストを発見することができる。

## 2.2 歌詞のトピックを活用した歌詞探索機能

LYRICS RADAR では、ユーザがトピックレーダー上で自ら選択して歌詞を見つける機能に加え、トピック比率をより一層活用した「トピック比率の直接入力機能」による検索機能を提案する。

### 2.2.1 トピック比率の直接入力機能：トピックレーダーチャートをクエリとした検索

トピックレーダーチャート自体を入力インターフェースとみなし、トピック比率の 5 個の値を図形として直接変形操作することでクエリとし、そのトピック比率に最も近い歌詞を検索できる機能である。各トピックの代表語を参考に、もっとこういったトピックを含む歌詞を検索したい、という検索要求に対応できる。図 3 に、五角形の左上のトピックが突出していた歌詞から出発して、五角形の真上のトピック比率をマウスのドラッグ操作で増加させた具体例を示す。このドラッグ操作中に、その歌詞の内容とトピックレーダー上での位置をリアルタイムに更新することで、ユーザは興味のある歌詞が表示された時点で操作をやめることができ、より直感的で探索的な歌詞検索が実現できた。

## 3. LYRICS RADAR の実装

LYRICS RADAR の中核となる歌詞のトピック分析は、代表的なトピックモデルである潜在的ディリクレ配分法 (Latent Dirichlet Allocation: LDA)[12] を用いて実現した。LDA では、歌詞を構成する各単語を異なるトピックに割り当てていくため、歌詞全体では複数のトピックから構成されているとみなすことができる。我々の目的では、多数の楽曲の歌詞を一度に与えて、それらを構成する代表

的な  $K$  個のトピックと各歌詞におけるトピックの比率を教師なしで推定する。その結果から、トピックごとに各単語の出現確率が決まるので、トピックを表す代表的な単語 (代表語) を求めることもできる。

### 3.1 歌詞に対する Latent Dirichlet Allocation (LDA)

LDA におけるモデル学習用のデータとして  $D$  個の独立した歌詞  $\mathbf{X} = \{\mathbf{X}_1, \dots, \mathbf{X}_D\}$  を考える。歌詞  $\mathbf{X}_d$  は、 $N_d$  個の単語系列  $\mathbf{X}_d = \{\mathbf{x}_{d,1}, \dots, \mathbf{x}_{d,N_d}\}$  で構成されている。歌詞に登場する全ての語彙の数を  $V$  とすると、 $\mathbf{x}_{d,n}$  は、語彙の中から選ばれた単語に対応する次元のみが 1 で、他は 0 である  $V$  次元ベクトルで表せる。また、文書  $\mathbf{X}_d$  に対応する潜在変数系列 (トピック系列) を  $\mathbf{Z}_d = \{\mathbf{z}_{d,1}, \dots, \mathbf{z}_{d,N_d}\}$  とする。トピック数を  $K$  とすると、 $\mathbf{z}_{d,n}$  は選ばれたトピックに対応する次元のみが 1 で他は 0 である  $K$  次元のベクトルで表せる。ここで、全ての歌詞の潜在変数系列をまとめて  $\mathbf{Z} = \{\mathbf{Z}_1, \dots, \mathbf{Z}_D\}$  とする。このときグラフィカルモデル (図 4) から変数間の条件付き独立性を考慮すると、完全な同時分布は

$$p(\mathbf{X}, \mathbf{Z}, \boldsymbol{\pi}, \boldsymbol{\phi}) = p(\mathbf{X}|\mathbf{Z}, \boldsymbol{\phi})p(\mathbf{Z}|\boldsymbol{\pi})p(\boldsymbol{\pi})p(\boldsymbol{\phi}) \quad (1)$$

で与えられる。ここで、 $\boldsymbol{\pi}$  は各歌詞におけるトピックの混合比 ( $D$  個の  $K$  次元ベクトル)、 $\boldsymbol{\phi}$  は各トピックにおける語彙の登場確率 ( $K$  個の  $V$  次元ベクトル) である。最初の二項には多項分布に基づく離散分布を仮定する。

$$p(\mathbf{X}|\mathbf{Z}, \boldsymbol{\phi}) = \prod_{d=1}^D \prod_{n=1}^{N_d} \prod_{v=1}^V \left( \prod_{k=1}^K \phi_{k,v}^{z_{d,n,k}} \right)^{x_{d,n,v}} \quad (2)$$

$$p(\mathbf{Z}|\boldsymbol{\pi}) = \prod_{d=1}^D \prod_{n=1}^{N_d} \prod_{k=1}^K \pi_{d,k}^{z_{d,n,k}} \quad (3)$$

残りの二項には多項分布の共役事前分布であるディリクレ分布を仮定する。

$$p(\boldsymbol{\pi}) = \prod_{d=1}^D \text{Dir}(\boldsymbol{\pi}_d|\boldsymbol{\alpha}^{(0)}) = \prod_{d=1}^D C(\boldsymbol{\alpha}^{(0)}) \prod_{k=1}^K \pi_{d,k}^{\alpha_{d,k}^{(0)}-1} \quad (4)$$

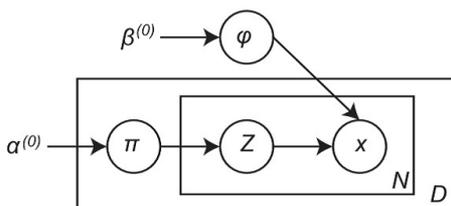


図 4 LDA のグラフィカルモデル

$$p(\phi) = \prod_{k=1}^K \text{Dir}(\phi_k | \beta^{(0)}) = \prod_{k=1}^K C(\beta^{(0)}) \prod_{v=1}^V \phi_{k,v}^{\beta_{k,v}^{(0)} - 1} \quad (5)$$

ここで、 $\alpha^{(0)}$  および  $\beta^{(0)}$  はハイパーパラメータである。 $C(\alpha^{(0)})$  および  $C(\beta^{(0)})$  は正規化定数であり、

$$C(\mathbf{x}) = \frac{\Gamma(\hat{x})}{\Gamma(x_1) \cdots \Gamma(x_I)}, \quad \hat{x} = \sum_{i=1}^I x_i \quad (6)$$

である。

各歌詞のトピック混合比である  $\pi$  は、正規化することでトピックレーダーチャートとして用いた。また、各トピックにおける語彙の出現確率  $\phi$  は、トピックレーダーチャートの各トピックと関連の高い代表語を求めるために用いた。

### 3.2 LDA の学習

LDA の学習における歌詞データとして、日本語歌詞のポピュラー音楽 (J-POP) 21845 曲から、その歌詞に 100 語以上が含まれていた 6902 曲を選別して用いた。この 6902 曲のアーティスト数は 1845 組で、作詞家の数は 2285 人であった。また、アーティスト単位でのトピック分析は、各アーティストが持つ単語を足しあわせた文章において 100 語以上が含まれていた 2848 アーティストを選別して用いた。語彙の数  $V$  は、データベース中のすべての歌詞において 10 回以上使用された 26229 語彙を用いた。歌詞の形態素解析には MeCab[13] を使用し、名詞、動詞、形容詞を抽出してその原形を一単語として数えた。ただし、複数の歌詞に幅広く登場しすぎる単語は一般的過ぎて、トピック分析を適切に行う上で支障がある。そこで、トピック分析時にそのような単語の重要度を下げたため、各単語に対して idf (Inverse Document Frequency) による重み付けを行った。

LDA の学習においては、トピック数を前述したとおり  $K = 5$  として学習した。ハイパーパラメータ  $\alpha^{(0)}$  と  $\beta^{(0)}$  の初期値はすべて 1 とした。

### 3.3 トピック数の評価

言語モデルの性能評価に用いられるパープレキシティを各トピック数においての算出した。図 5 はパープレキシティの算出結果である。赤色の点は各歌詞に、青色の点は各アーティストに LDA を適用した場合のパープレキシティを表す。パープレキシティは値が高いほどモデルが複

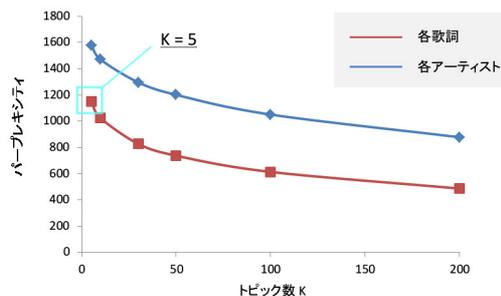


図 5 各トピック数におけるパープレキシティの値

雑であることを示す。よって、パープレキシティの値が低いほど言語モデルとして性能が高いと評価できる。この際、パープレキシティの算出には以下の式を用いた。

$$\text{perplexity}(\mathbf{X}) = \exp\left(-\frac{\sum_{d=1}^D \log p(X_d)}{\sum_{d=1}^D N_d}\right) \quad (7)$$

結果から、現在の実装で用いたトピック数 ( $K = 5$ ) では、1150 であった。図 5 からは、トピック数を増やすほどパープレキシティが低い値になっていることがわかるが、2.1.1 で前述したようにトピック数と操作性のトレードオフが存在するため、それらのバランスを見極めた上で、トピック数の増加やユーザインタフェースの改良などに関し今後の研究の余地がある。

## 4. 評価実験

LYRICS RADAR におけるトピック分析結果 (トピックレーダーチャートや歌詞のマッピングに関係) の妥当性を検証するため、主観評価実験を行った。被験者は 20 代男性 17 名、女性 4 名の計 21 名で、3.2 節で述べた 6902 曲の歌詞を対象にした LDA の結果を用いた。

### 4.1 トピック分析の評価

LDA によるトピック分析の結果、推定されたトピック比率がその歌詞を適切に表現できていることを検証した。さらに、ある歌詞を指定した際、その歌詞と同一のアーティストや同一の作詞家による歌詞が、トピックとしてお互いにどれだけ近いのかも同時に調査した。

#### 4.1.1 実験方法

ランダムに選出した歌詞 1 曲 (以降、選出歌詞と呼ぶ) に対し、それを基準に選ばれた下記の 4 曲の歌詞 (以降、比較歌詞と呼ぶ) と比較した。

- (1) 歌詞マップ上で、選出歌詞と最も距離の近い歌詞
- (2) 選出歌詞と同じ作詞家の歌詞の中で、最も距離の近い歌詞
- (3) 選出歌詞と同じアーティストの歌詞の中で、最も距離の近い歌詞
- (4) ランダム選出による歌詞

	タイトル	アーティスト	作詞家
選出された歌詞	雪の華	中島美嘉	Satomi
(A1) 最も距離の近い歌詞	December	I WISH	ai
(A2) 同じ作詞家	心に夢を君には愛を	Kinki kids	Satomi
(A3) 同じアーティスト	TEARS(粉雪が舞うように...)	中島美嘉	秋元康・中島美嘉
(A4) ランダム	まゆげ	東京プリン	伊藤洋介

図 6 評価実験（トピック分析の評価）において使用した歌詞

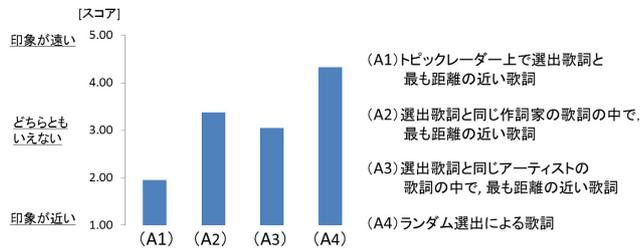


図 7 評価実験（トピック分析の評価）の結果. (1) のスコアが最も 1.0 に近く、適切に分析できていることがわかった.

これら 4 曲の歌詞の情報を図 6 に示す.

被験者は、選出歌詞を見た後、選出歌詞と比較歌詞の一つを見比べながら、二つの歌詞から受ける印象の近さを 5 段階 (1: 近い, 2: なんとなく近い, 3: どちらともいえない, 4: なんとなく遠い, 5: 遠い) で評価した. ここで、比較歌詞 (1)~(4) の被験者への提示順序はランダムとした. また各比較歌詞の評価が終わった段階で、なぜその評価にしたのかという基準を可能な範囲で記述して回答させた.

#### 4.1.2 実験結果

実験結果を図 7 に示す. 4 曲の比較歌詞において、全被験者による 5 段階評価結果の平均スコアがグラフ表示されている. 比較歌詞 (1) の結果が最も 1 に近かったことから、歌詞マップ上で最も距離の近い歌詞が、選出歌詞の印象と最も近かったと適切に判断されていたことがわかる. 比較歌詞 (2) と (3) の結果は 3 に近く、同じアーティストや同じ作詞家の歌詞の中で最も距離の近い歌詞は、ほぼ「どちらともいえない」と判断されていた. また、ランダムで選出した比較歌詞 (4) は、印象が遠いと適切に判断されていた.

なぜその評価にしたのかという評価基準に関する質問では、季節感、ネガティブ、ポジティブ、恋愛、人間関係、色、明暗情報、自分主体か否か、ノリ、テンションといった回答結果が得られた.

#### 4.1.3 考察

選出歌詞の楽曲数を増やした詳細な評価実験をしなければ最終的な結論は得られないが、今回の選出歌詞 1 曲の実験結果では、少なくとも歌詞マップ上での距離が最も近い歌詞が最も近い印象を持っていると評価され、ランダム選出した歌詞の印象が遠かったことから、トピック分析がある程度有効に機能していることを確認できた.

評価基準の回答では、人によって基準が大きく異なる場合があり、例えば、歌詞の季節感が似ているだけで似た印象を感じる人もいれば、同じ季節であっても、それ以外の

要素 (例えば歌詞中の登場人物の心情など) に違いがあると似た印象を受けない人がいることがわかった.

## 5. おわりに

本研究では、ユーザが好みの歌詞を検索したり、新たな歌詞に出会ったりするために、歌詞の潜在的なトピックの比率に基づいてインタラクティブに歌詞を検索できるインタフェース LYRICS RADAR を提案した. 本研究の学術的な意義は、従来表層的にしか扱うことが困難だった歌詞に対し、その深層的な意味をトピックレーダーチャートとして表現することで、トピックの可視化とインタラクティブで多様な入力手段を共に可能にしたことにある. ユーザにトピックレーダーチャートの五角形の形状を直接操作させることで、トピックに基づく歌詞の検索を実現した. そして、トピックレーダーチャートの形状に近い歌詞を近くに配置した二次元のマッピングにより、検索クエリを思い浮かなくても様々な歌詞をブラウジングしながら、新たな歌詞に出会うことを可能にした.

今後の研究としては、個々のユーザの違いを反映したユーザ適応型インタフェースの実現や、階層的なトピック分析 [14] によるトピック分析機能の高度化等が考えられる. また、歌詞の持つ細かなトピックに対応した検索インタフェースを実現するために、より多くのトピック数を反映できるような可視化手法を考えることが今後の検討課題である.

謝辞 本研究の一部は JST CREST 「OngaCREST プロジェクト」の支援を受けた.

## 参考文献

- [1] 森 数馬: 日常の音楽聴取における歌詞の役割についての研究, 対人社会心理学研究, 10, pp. 131-137, (2010).
- [2] Müller, M., Kurth, F., Damm, D., Fremerey, C., and Clausen, M.: Lyrics-based Audio Retrieval and Multimodal Navigation in Music Collections, Proc. of ECDL, pp. 112-123, (2007).
- [3] Neumayer, R. and Rauber, A.: Multi-modal Music Information Retrieval: Visualisation and Evaluation of Clusterings by Both Audio and Lyrics, Proc. of RIAO, pp. 70-89, (2007).
- [4] Laurier, C., Grivolla, J., and Herrera, P.: Multimodal Music Mood Classification Using Audio and Lyrics, Proc. of ICMLA, pp. 688-693, (2008).
- [5] Neumayer, R. and Rauber, A.: Integration of Text and Audio Features for Genre Classification in Music Information Retrieval, Proc. of ECIR, pp. 724-727, (2008).
- [6] Brochu, E. and de Freitas, N.: "Name That Song!": A Probabilistic Approach to Querying on Music and Text, Proc. of NIPS, (2002).
- [7] Takahashi, R., Ohishi, Y., Kitaoka, N., and Takeda, K.: Building and Combining Document and Music Spaces for Music Query-By-Webpage System, Proc. of Interspeech 2008, pp. 2020-2023, (2008).
- [8] Fujihara, H., Goto, M., Ogata, J., and Okuno, H.G.: LyricSynchronizer: Automatic Synchronization System

- between Musical Audio Signals and Lyrics, IEEE Journal of Selected Topics in Signal Processing, 5(6), pp. 1252–1261, (2011).
- [9] 中野 倫靖, 後藤 真孝: VocaRefiner: 歌を歌って歌い直して統合できる新しい歌声生成インタフェース, WISS 2012 論文集, pp. 1–6, (2012).
- [10] 都築 圭太, 中野 倫靖, 後藤 真孝, 山田 武志, 牧野 昭二: 様々な歌手が同じ曲を歌った歌声の多様性を活用するシステム, 情報処理学会 音楽情報科学研究会 研究報告, 100(21), pp. 1–8, (2013).
- [11] Maaten, L. and Hinton, G. E.: Visualizing data using t-SNE, Journal of Machine Learning Research, 9, pp. 2579–2605, (2008).
- [12] Blei, D. M., Ng, A. Y., and Jordan, M. I.: Latent Dirichlet Allocation, Journal of Machine Learning Research, 3, pp. 993–1022, (2003).
- [13] 工藤 拓: MeCab: Yet Another Part-of-Speech and Morphological Analyzer, <http://mecab.googlecode.com/svn/trunk/mecab/doc/index.html>.
- [14] Adams, R., Ghahramani, Z., and Jordan, M.: Tree-Structured Stick Breaking Processes for Hierarchical Data, Proc. of NIPS, 23, pp. 19–27, (2010).