

LyricsRadar：歌詞の潜在的意味に基づく 歌詞検索インタフェース

佐々木 将人^{1,a)} 吉井 和佳^{2,b)} 中野 倫靖^{3,c)} 後藤 真孝^{3,d)} 森島 繁生^{1,e)}

受付日 2015年7月29日, 採録日 2016年2月8日

概要：本論文では、歌詞の潜在的意味に基づいた検索インタフェース *LyricsRadar* 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、まった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、ユーザー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、本研究では潜在的ディリクレ配分法を用いて歌詞が潜在的に持つ意味（トピック）を自動的に分析し、歌詞のトピックの可視化を行った。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および、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した歌詞マップという2つの可視化が特徴である。6,902曲の歌詞データを用いたトピック分析結果の妥当性の評価および、既存検索システムとの比較による被験者実験により、*LyricsRadar* が有効に機能していることを示した。

キーワード：歌声情報処理, ユーザインタフェース, 歌詞検索, 可視化, Latent Dirichlet allocation

LyricsRadar: A Lyrics Retrieval Interface Based on Latent Topics of Lyrics

SHOTOKO SASAKI^{1,a)} KAZUYOSHI YOSHII^{2,b)} TOMOYASU NAKANO^{3,c)} MASATAKA GOTO^{3,d)}
SHIGEO MORISHIYAMA^{1,e)}

Received: July 29, 2015, Accepted: February 8, 2016

Abstract: This paper presents a lyrics retrieval interface called *LyricsRadar* that enables users to interactively browse song lyrics by visualizing their topics. Since conventional lyrics retrieval systems are based on simple word search, those systems often fail to reflect user's intention behind a query when a word given as a query can be used in different contexts. For example, the word “tears” can appear not only in sad songs (*e.g.*, feel heartrending), but also in happy songs (*e.g.*, weep for joy). To overcome this limitation, we propose to automatically analyze and visualize topics of lyrics by using a well-known text analysis method called latent Dirichlet allocation (LDA). This enables *LyricsRadar* to offer two types of topic visualization. One is the topic radar chart that visualizes the relative weights of five latent topics of each song on a pentagon-shaped chart. The other is radar-like arrangement of all songs in a two-dimensional space in which song lyrics having similar topics are arranged close to each other. The subjective experiments using lyrics of 6,902 Japanese popular songs showed that *LyricsRadar* can appropriately navigate users to lyrics of interests.

Keywords: singing information processing, user interface, lyrics retrieval, visualization, latent Dirichlet allocation

¹ 早稲田大学
Waseda University, Shinjuku, Tokyo 101-0062, Japan

² 京都大学
Kyoto University, Kyoto 606-8501, Japan

³ 産業技術総合研究所
National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Ibaraki 305-8568, Japan

a) joudanjanai-ss@akane.waseda.jp

b) yoshii@kuis.kyoto-u.ac.jp

c) t.nakano@aist.go.jp

d) m.goto@aist.go.jp

e) shigeo@waseda.jp

1. はじめに

歌唱をともなうポピュラー音楽では歌詞は重要な要素の1つであり [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分からないことが多い. 歌詞中に出現する語句を表層的にテキスト全文検索することだけでは, 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点で, 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検索インタフェースを実現することを目的とする.

本論文で対象とする歌詞の潜在的意味 (トピック) は機械学習手法 (潜在的ディリクレ配分法) によって推定されるが, それは歌詞を構成する単語のトピック比率の足し合わせとして定義される. すなわち, ある歌詞のトピックは, そこにどのような単語が含まれているかによって決まる. たとえば, 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」等の代表語を持つトピックと「桜, 花びら, 吹雪, 葉」等の代表語を持つトピックの2つに属していた. 代表語から推測すると, 前者は「人が踊る動作」というトピックにおける「舞う」, 後者は「花びらや雪等の軽い物体の動き」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.

従来, 歌詞に基づいた楽曲検索や分類に関して, 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等で分類, マッピングを行う研究 [2], [3], [4], [5] がなされてきた. また, 歌詞の自動トピック推定 [6] や意味の解析 [7] を行った研究もある. 歌詞の潜在的な意味を用いて楽曲の楽譜とその歌詞との関係をモデル化し, 音符列と歌詞に基づく楽曲検索を可能にする研究 [8], [9] 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 [10] 等もあった. 他にも, 楽曲の可視化のために歌詞を用いた研究 [11] や, 歌詞にあったスライドショーを生成する研究 [12] 等があげられる. これに対して本論文では, インタラクティブな歌詞検索インタフェースを提案する点で異なる. 本提案は, 歌詞のみを対象としている点から, 特に歌詞の解析やマッピングに関する研究 [2], [3], [4], [5], [6], [7] や, 歌詞に加えて音符列も考慮する研究 [8], [9] と関係が深い. しかし, 歌詞のような個々人の嗜好の違いがある検索対象においては, 好きな単語や好みの歌詞をクエリとして検索するだけでなく, 検索結果を受けた先のインタラクションを含めた設計が必要であると考えており, その点で新しい.

ユーザが自分好みの歌詞を検索する既存の技術として

は, 既存の歌詞検索サイトがある. このようなサイトでは, 「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があるが, 人手でのタグ付けには限界があるだけでなく, 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. また, ランキングは効率的な検索手段の1つであるが, Macrae らは, オンライン検索において歌詞のランキングの精度が低いことを示し, 高精度な歌詞のランキング手法について提案した [13]. さらに, 検索以外に歌詞を活用した音楽インタフェースとして, 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 [14] や, 歌唱の録音において歌詞を活用する VocaRefiner [15] 等があった. しかしこれらは, 好みの歌詞を探すために歌詞の潜在的なトピックを用いていなかった.

そこで本研究では, 歌詞が潜在的に持つ意味 (トピック) を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「LyricsRadar」を提案する. 具体的には, 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とで, 似たトピックを持つ歌詞を検索することを可能とする. この 5 次元ベクトルで表される歌詞のトピック比率を, 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分かり, 直感的な歌詞検索・可視化が可能になる. また, トピックレーダーチャートを直接変形させることで歌詞を検索することもできる.

LyricsRadar 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) が近くになるように, すべての歌詞を二次元平面上に配置する. これは, 5 次元のトピック比率を 2 次元に次元圧縮することで実現した (2.1.2 項で後述). ユーザは, この平面上を探索することで, 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. ある歌詞から見て, 距離は同じでも方向 (位置) が違う複数の歌詞については, トピックの異なり方に違いがある. ここで, どういった違いかについては, トピックレーダーチャートで確認できる. 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. LyricsRadar の機能

LyricsRadar は, 各楽曲の歌詞の持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フェースである. そのために, 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



図 1 LyricsRadar 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)

Fig. 1 An example display of the lyrics retrieval interface of LyricsRadar. The lyrics in this screenshot are taken from the RWC Music Database (RWC-MDB-P-2001 No.30), although LyricsRadar is implemented using Japanese popular songs (J-POP).

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分からない曖昧な状況でも活用できるという特長を持つ。

また、LyricsRadar の機能には、既存の Web 歌詞検索システムで一般的に用いられている、歌手名・曲名の一覧から選択する機能や、歌手名・曲名・歌詞をクエリ入力により検索する機能も含まれている。

2.1 歌詞のトピックに基づく可視化機能

LyricsRadar 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。以降、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

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る (図 1 左上)。五角形の各頂点がそれぞれ異なるトピックに対応し、各トピックと関連の高い代表語が五角形の外側に 3 つ表示される (各トピックにおける単語の出現確率

が上位の 3 単語)。このようなトピックレーダーチャートによって、ユーザは歌詞の (表層的な) 単語以外の観点として、各トピックの意味や個々の歌詞の内容を把握することができる。

トピックレーダーチャートは外周に近いほど大きな比率を表す。また、トピック比率を可視化したものであるため、5 つの次元の値の合計はつねに一定値である。たとえば、ある 1 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。この特徴により、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 と決定した。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。トピックを学習する際に用いた歌詞データベースを変更した場合には、適切なトピックを再度推定し直すことが望ましい。その際には、データベースに含まれる歌詞によってトピックレーダーチャートの意味や、表示される代表語が動的に変わることになる (トピック数は 5 で固定)。

2.1.2 歌詞をマッピングした二次元平面

マッピングは、データベース中のすべての歌詞をそのト

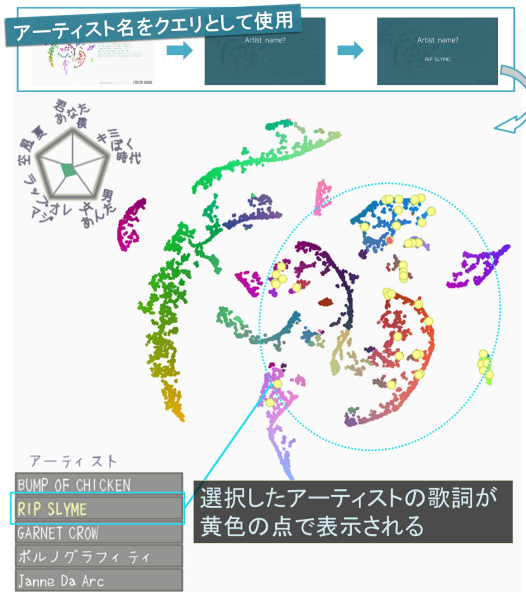


図 2 アーティスト名選択による歌詞の表示例

Fig. 2 An example display of selecting the lyrics of a single artist.

ピックの類似度に応じて二次元平面上に配置し、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る(図 1 左). 二次元平面上へのマッピングには Maaten らが提案した t-SNE [16] を用い、各歌詞のトピック比率 (5 次元) が似ているほど配置が近くなるように次元圧縮した. マッピングの結果は、その一部を拡大することができる.

さらに、各歌詞に対応する点はトピック比率に応じて着色することで、2次元に圧縮してマッピングした際に失われた情報を色 (3次元) の違いで表現する. 色付けにおける次元圧縮においても、同様に t-SNE [16] を用いた. 5次元のトピック比率を三次元に圧縮した後、各次元を正規化して RGB に対応付けた. 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.

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、作詞家ごとによりに分布しているかを直

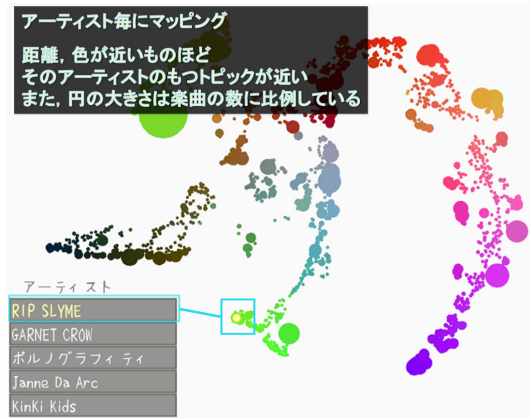


図 3 ポピュラー音楽 (J-POP) の 2,848 組のアーティストのマッピング

Fig. 3 Mapping of 2,848 artists in J-POP according to their lyrics.

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ても、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、同一アーティストの複数歌詞をまとめて1つの長い歌詞と見なし、それらを用いて新たにトピック分析することで、アーティスト単位でのマッピングも可能である. 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 3 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

LyricsRadar 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いっそう活用した「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フェースと見なし、5次元ベクトルで表されるトピック比率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 4 に、五角形の右上のトピックが突出していた歌詞のトピックレーダーチャートに基づいて、それを変形させた場合の具体例を示す. 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. このドラッグ操作中に

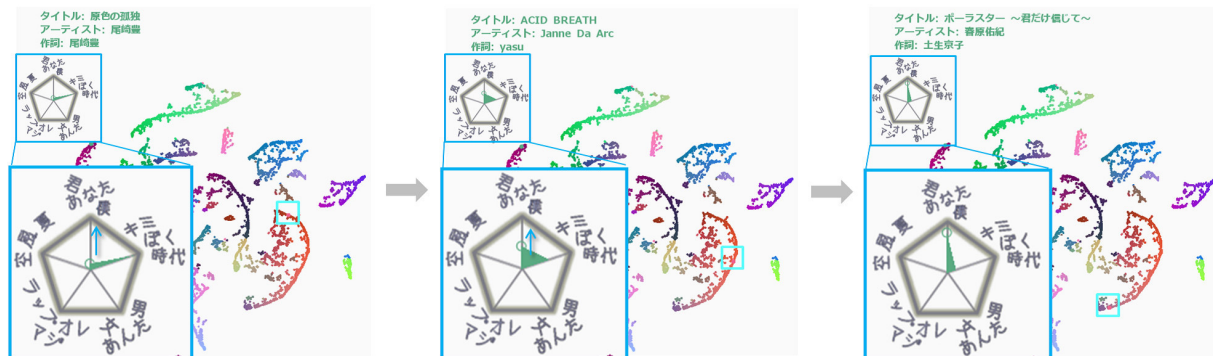


図 4 トピックレーダーチャート上でトピック比率を直接入力して歌詞を検索する具体例：五角形の上部の比率をマウスのドラッグにより上へ伸ばした様子を示す

Fig. 4 An example of the direct manipulation of the topic ratio on the topic radar chart: illustration of a user changing the ratio of the top-most dimension in the pentagon-shaped chart.

は、その歌詞の内容とトピックレーダー上での位置がリアルタイムに更新されるので、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等、より直感的で探索的な歌詞検索が実現できる。

3. LyricsRadar の実装

LyricsRadar 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) [17] 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、1つの歌詞を複数のトピックから構成されていると見なすことができる。

本研究では、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。文書 \mathbf{X}_d は、 N_d 個の単語系列 $\mathbf{x}_d = \{x_{d,1}, \dots, x_{d,N_d}\}$ で構成されている。3.2 節で後述するが、本論文では1つの文書 \mathbf{X}_d として、1つの歌詞、もしくは同一アーティストのすべての歌詞 (名詞、動詞、形容詞のみ) に対応する。歌詞データベース中で出現するすべての語彙の数を V とすると、 $x_{d,n}$ 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 $\mathbf{Z}_d = \{z_{d,1}, \dots, z_{d,N_d}\}$ とする。トピック数を K とすると、 $z_{d,n}$ 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、すべ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

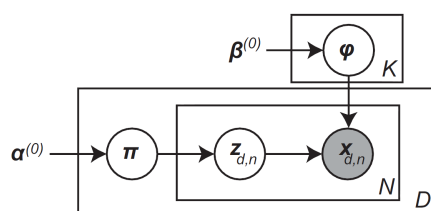


図 5 LDA のグラフィカルモデル

Fig. 5 Graphical representation of latent Dirichlet allocation (LDA).

デル (図 5) 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出現確率 (K 個の V 次元ベクトル) である。最初の2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2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。

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

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

ここで、 $\boldsymbol{\alpha}^{(0)}$ および $\boldsymbol{\beta}^{(0)}$ はハイパーパラメータである。 $C(\boldsymbol{\alpha}^{(0)})$ および $C(\boldsymbol{\beta}^{(0)})$ 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 π は、ディリクレ事後分布のパラメータ期待値を求め、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 ϕ も同様に期待値を求め、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

歌詞のトピックを正確に推定するために、歌詞の選別およ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、1 楽曲の中で用いることのできる単語の数は限られている。さらに、A メロやサビ等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出現する単語数は少ない。トピック分析は、文章内の単語に依存するため、単に歌詞を文章と見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、語彙数に閾値を設定して歌詞を選別したり、同一アーティストの複数歌詞をまとめて扱うことで語彙数を増やしたりしたうえでトピック分析することで、歌詞の持つトピックの比率を求めた。

具体的には、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) 21,845 曲から、その歌詞に 100 語彙以上が含まれていた 6,902 曲を選別して用いた。この 6,902 曲のアーティスト数は 1,845 組で、作詞家の数は 2,285 人であった。また、アーティスト単位でのトピック分析では、それぞれアーティストが持つ歌詞をつなぎ合わせた文章において、100 語彙以上が含まれていた 2,848 アーティストを選別して用いた。語彙の総数 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 26,229 語彙を用いた。

歌詞の形態素解析には MeCab [18] を使用し、名詞、動詞、形容詞を抽出してその原形を 1 単語として数えた。ただし、複数の歌詞に幅広く出現する単語は一般的すぎて、トピック分析を適切に行ううえで支障がある。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るため、各単語に対して重み付けを行った。具体的には、

$$idf_i = \log \frac{D}{df_i} \quad (7)$$

表 1 評価実験 (トピック分析の評価) において使用した歌詞: 選出歌詞と歌詞マップ上で最も近い比較歌詞 (A1)

Table 1 List of lyrics used in the evaluation experiment (evaluation of topic analysis): basis lyrics (left) and the lyrics closest to them (right) on the lyrics map.

選出歌詞	比較歌詞 (A1)
HEAT CAPACITY (T.M.Revolution) かつおぶし (さくらと一郎)	優しくて少しバカ (嵐) 道頓堀情話 (弦哲也・綾世一美)
RESPECT the POWER OF LOVE (安室奈美恵)	愛したひとはバツイチ (ナインティナイン&清水ミチコ)
ACT ON MY STYLE (MCU feat.RATHER UNIQUE&川上次郎)	フラストレーションミュージック (Hysteric Blue)
HOW TO LOVE (広瀬香美)	ブギウギ時代 (中村美津子)

のように求めた idf (Inverse Document Frequency) を、各歌詞中の各単語に対する各トピックの負担率にかけ合わせることで重み付けを行った。 D は総歌詞数、 df_i は単語 i が出現する歌詞の数を表す。

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。

4. 評価実験

LyricsRadar 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) の妥当性および、ユーザビリティを評価するため、2 つの被験者実験による評価を行った。本章では、それぞれの実験の方法と結果を述べて考察する。それぞれの実験は、ともに 3.2 節で述べた 6,902 曲の歌詞を対象にした LDA の結果を用いた。

4.1 被験者実験による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評価した。被験者は 20 代男性 3 人、30 代男性 3 人、40 代男性 1 人、30 代女性 1 人の計 8 人である。

4.1.1 実験方法

ランダムに選出した歌詞 5 曲を入力 (選出歌詞と呼ぶ) とし、それを基準に選ばれた下記の 5 曲の歌詞 (比較歌詞と呼ぶ) と比較した。

- (歌詞 A1) 選出歌詞と最も距離の近い歌詞
- (歌詞 A2) 選出歌詞からの距離が 1/4 分位となる歌詞
- (歌詞 A3) 選出歌詞からの距離が中央値となる歌詞
- (歌詞 A4) 選出歌詞からの距離が 3/4 分位となる歌詞
- (歌詞 A5) 選出歌詞と最も距離の遠い歌詞

ここで、選出歌詞と比較歌詞の距離は、歌詞マップ (二次元平面) 上の座標のユークリッド距離とした。選出歌詞とそれ以外のすべての歌詞との距離を計算した後、距離が最小となる歌詞、距離の 1/4 分位数に最も近い歌詞、距離の中央値に最も近い歌詞、距離の 3/4 分位数に最も近い歌詞、距離が最小となる歌詞を選んだ。選出歌詞と歌詞 A1 に関する情報を表 1 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 1

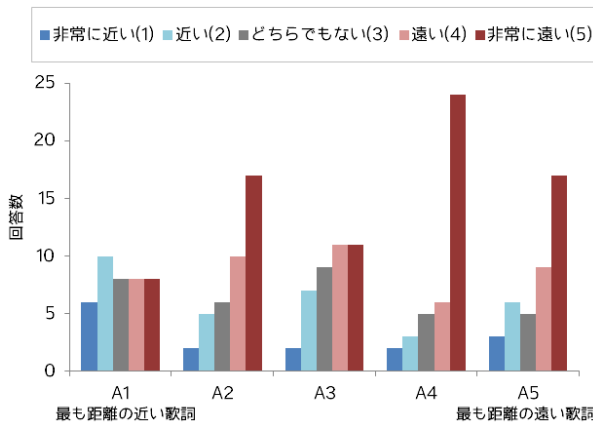


図 6 評価実験 (トピック分析の評価) の結果。被験者 8 人分の評価の回答数を示す

Fig. 6 Results of the evaluation experiment (evaluation of topic analysis). This figure shows the number of answers by eight subjects.

つとを見比べながら、2つの歌詞から受ける印象の近さを5段階 (1:非常に近い, 2:近い, 3:どちらともいえない, 4:遠い, 5:非常に遠い) で評価した。ここで、比較歌詞 (1)~(5) の被験者への呈示順序はランダムとしたが、評価の修正に関して制限は設けなかった。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 6 に示す。5 曲の歌詞 A1~A5 において、全被験者による 5 段階評価結果の回答数がグラフ表示されている。A1 の結果が最も 1 (非常に近い) および 2 (非常に近い) の回答数が多く、5 (非常に遠い) が少なかった。したがって、歌詞マップ上で最も距離の近い歌詞は、入力の影響と近く判断される傾向にあった。それ以外の比較歌詞 (A2~A4) については、今回の被験者間では、回答に違いが見られなかったが、選出歌詞を「ACT ON MY STYLE (MCU feat.RATHER UNIQUE&川上次郎)」とした実験に関しては、比較歌詞の距離が近いほど被験者の評価が 1 に近づく傾向があった (図 7)。

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、テーマ (恋愛、人生、未来等)、主人公 (キャラ)、デュエットかどうか、時代、ポジティブ/ネガティブ、情熱、長さ、口調、英語・カタカナの比率、自分中心か相手中心か、構成、音楽ジャンルの違い、といった回答が得られた。

4.1.3 考察

歌詞 5 曲を入力として用いた今回の実験結果からは、歌詞マップ上で距離が最も近い歌詞は、距離が遠い歌詞と比較して類似していると評価されることが多かった。したがって、「歌詞マップ上で近い歌詞の印象は近い」というインタフェース性能について確認できた。

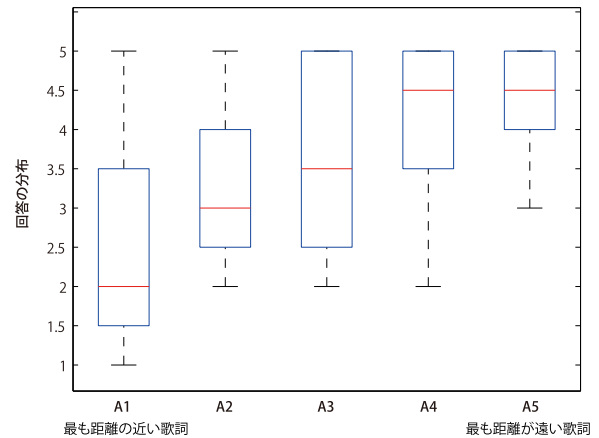


図 7 選出歌詞を「ACT ON MY STYLE (MCU feat.RATHER UNIQUE&川上次郎)」とした実験結果

Fig. 7 Results when “ACT ON MY STYLE (MCU feat.RATHER UNIQUE&Jiro Kawakami)” is used as the basis lyrics.

評価基準の回答では、人によって多様な基準があることを示した。たとえば、季節感やテーマといった歌詞の意味的な情報を重視している人もいれば、カタカナや英語を使うかどうか、構成、言葉遣いといった表現方法に着目する人がいることが分かった。

4.2 被験者実験によるユーザビリティの評価

LyricsRadar のユーザビリティの評価を行った。国際規格である ISO 規格・ISO9241-11 において、ユーザビリティ (Usability) は『ある製品が指定された利用者によって、指定された利用の状況下で、指定された目的を達成するために用いられる際の有効さ、効率および利用者の満足度の度合い』と定義づけられている。この定義に基づき、LyricsRadar の効率・有効性・満足度について評価した。

4.2.1 実験方法

被験者には、未知かつユーザ好みの歌詞を探索することを実験目的として提示した。効率および有効性の評価のため、どの程度自分好みの歌詞に出会えたか (4 段階)、目標歌詞発見までにどの程度時間がかかったかを既存歌詞検索インタフェースと比較した。また、満足度の評価のため、LyricsRadar やその機能に対してのアンケート (7 段階) を行った。

本実験で使用した既存歌詞検索インタフェースを図 8 に示す。比較のための既存検索機能として、Web 歌詞検索システムで一般的に用いられている「歌手名・曲名の一覧から選択する機能」および「歌手名・曲名・歌詞をクエリ入力により検索する機能」を実装した。

また、実験条件を統制して平等な状況で比較を行うために、既存検索インタフェースの歌詞データは LyricsRadar と同じ 6,902 曲の歌詞を使用し、インタフェースの使用順序は均等になるように被験者ごとに順序を変更した。被験

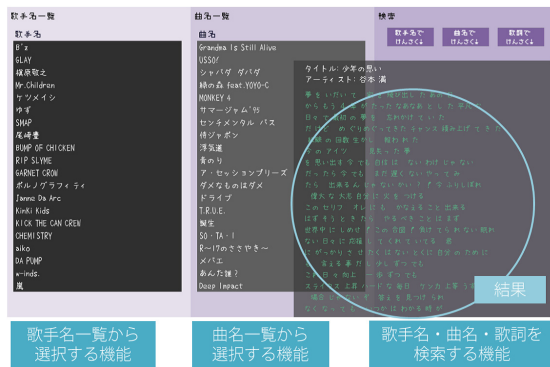


図 8 既存歌詞検索インタフェースの表示例(実際にはポピュラー音楽(J-POP)を用いて実装しているが、本図では歌詞の例示のためにRWC 研究用音楽データベースの楽曲(RWC-MDB-P-2001 No.30)を用いた)

Fig. 8 Examples of existing lyrics retrieval systems. The lyrics in this screenshot are taken from the RWC Music Database (RWC-MDB-P-2001 No.30), although *LyricsRadar* is implemented using Japanese popular songs (J-POP).

者は 20 代男性 8 人である。

実験の手順は、下記のとおりである。なお、被験者の半数は下記の手順で、もう半数は (4)–(6) の後に (0)–(3) を行った。

LyricsRadar による歌詞検索

(0) 操作説明

- (1) 未知かつユーザ好みの歌詞の探索
- (2) どれくらい好みの歌詞に出会えたか (4 段階), どれくらい時間がかかったかを回答
- (3) 歌詞マップおよびトピックレーダーチャートに対してのアンケート (7 段階) の回答および、どのように歌詞を検索したか等の自由記述

既存検索インタフェースによる歌詞検索

- (4) 操作説明
- (5) 未知かつユーザ好みの歌詞の探索
- (6) どれくらい好みの歌詞に出会えたか (4 段階), どれくらい時間がかかったかを回答

4.2.2 実験結果

実験の結果を図 9, 図 10 に示す。既存検索インタフェースと比較して, *LyricsRadar* を用いた場合, 好みの歌詞に出会うまで平均所要時間は約 2.5 分短く, 図 9 より, 出会えた歌詞の好みの度合いも良い結果となった。よって, 既存検索インタフェースと比べて, より有効な検索ができていることが確認できた。

また, 図 10 から歌詞マップ, トピックレーダーチャートともに有効に機能していることが確認でき, *LyricsRadar* に対してもまた使いたいという意見 (+1 以上) が 60% 以上を占めた。

LyricsRadar を用いた調べ方に関する記述回答では, 歌

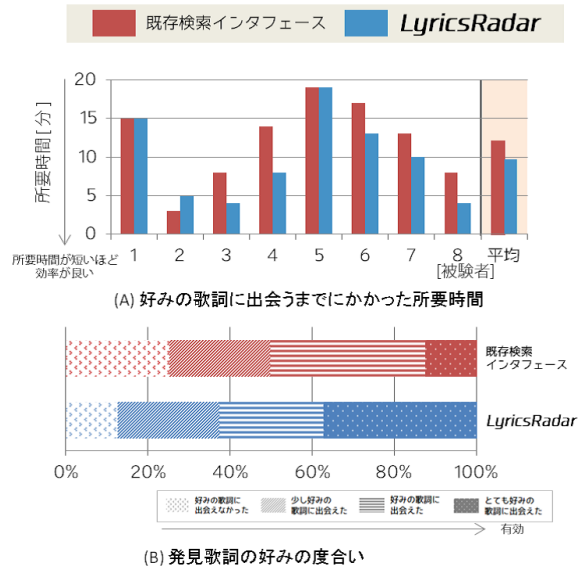


図 9 被験者実験による既存検索インタフェースとの比較結果
Fig. 9 Results of the subjective evaluation experiment based on comparisons with existing lyrics retrieval systems.

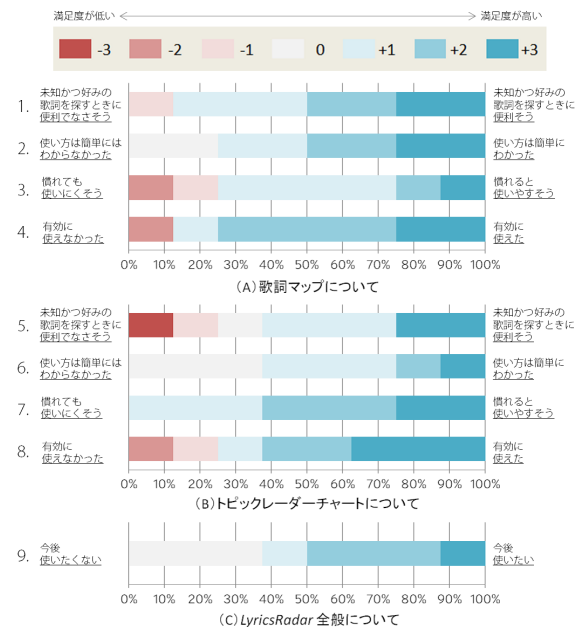


図 10 被験者実験による *LyricsRadar* とその機能に関するアンケートの結果
Fig. 10 Results of the subjective evaluation experiment based on the questionnaire for two functions of *LyricsRadar* and overall assessment.

詞クエリを入力し歌詞マップやトピックレーダーチャートで特徴をつかみながら調べた被験者や, トピックレーダーチャートをマウスでドラッグ操作しながら, それによって歌詞マップの特徴をつかむことに利用した被験者が多かった。また, あるアーティストに対して, 欲しい歌詞の雰囲気を出すことにトピックレーダーチャートを活用して歌詞を検索した被験者も存在した。その被験者の回答では, *LyricsRadar* を用いていると「ACIDMAN (アーティスト

名) + ‘愛’ = 横原敬之 (アーティスト名) の一部の歌詞」という法則があると感じ、好みのアーティストに雰囲気を感じながら検索することができた、と記述されていた。このように、トピックレーダーチャートは、「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する、という当初想定していた方法に加えて、他の形で活用もされていた。他に、検索だけでなく、歌詞の観点から楽曲やアーティストに対する理解が深まったという記述もあった。

5. おわりに

本論文では、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 *LyricsRadar* を提案した。本論文の学術的な意義は、歌詞を表層的な扱いのみでなく、その潜在的な意味をトピックレーダーチャートとして表現したうえで、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とも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。また被験者実験により、*LyricsRadar* の有用性をトピック分析の精度およびユーザビリティの観点から示した。

今後の展開としては、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 [20] 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や歌詞検索に最適な配置等が検討課題である。

謝辞 本論文の一部は、科学技術振興機構 OngaCREST プロジェクトによる支援を受けました。

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, No.10, pp.131–137 (2010).
 [2] Laurier, C., Grivolla, J. and Herrera, P.: Multimodal Music Mood Classification Using Audio and Lyrics, *Proc. ICMLA 2008*, pp.688–693 (2008).
 [3] McKay, C., Burgoyne, J.A., Hockman, J., Smith, J.B.L., Vigliensoni, G. and Fujinaga, I.: Evaluating the genre classification performance of lyrical features relative to audio, symbolic and cultural features, *Proc. ISMIR 2010*, pp.213–218 (2010).
 [4] Zaanen, M.V. and Kanters, P.: Automatic Mood Classification Using TF*IDF Based on Lyrics, *Proc. ISMIR 2010*, pp.75–80 (2010).

[5] Hu, Y., Chen, X. and Yang, D.: Lyric-based Song Emotion Detection with Affective Lexicon and Fuzzy Clustering Method, *Proc. ISMIR 2009*, pp.122–128 (2009).
 [6] Kleedorfer, F., Knees, P. and Pohle, T.: Oh Oh Oh Whoah! Towards Automatic Topic Detection In Song Lyrics, *Proc. ISMIR 2008*, pp.287–292 (2008).
 [7] Logan, B., Kositsky, A. and Moreno, P.: Semantic Analysis of Song Lyrics, *Proc. IEEE ICME 2004*, Vol.2, pp.827–830 (2004).
 [8] Brochu, E. and de Freitas, N.: “Name That Song!”: A Probabilistic Approach to Querying on Music and Text, *Proc. NIPS 2003*, pp.1505–1512 (2003).
 [9] Müller, M., Kurth, F., Damm, D., Fremerey, C. and Clausen, M.: Lyrics-based Audio Retrieval and Multimodal Navigation in Music Collections, *Proc. ECDL 2007*, pp.112–123 (2007).
 [10] Takahashi, R., Ohishi, Y., Kitaoka, N. and Takeda, K.: Building and Combining Document and Music Spaces for Music Query-By-Webpage System, *Proc. Interspeech 2008*, pp.2020–2023 (2008).
 [11] Neumayer, R. and Rauber, A.: Multi-modal Music Information Retrieval: Visualisation and Evaluation of Clusterings by Both Audio and Lyrics, *Proc. RAO 2007*, pp.70–89 (2007).
 [12] Funasawa, S., Ishizaki, H., Hoashi, K., Takishima, Y. and Katto, J.: Automated Music Slideshow Generation Using Web Images Based on Lyrics, *Proc. ISMIR 2010*, pp.63–68 (2010).
 [13] Macrae, R. and Dixon, S.: Ranking Lyrics for Online Search, *Proc. ISMIR 2012*, pp.361–366 (2012).
 [14] 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*, Vol.5, No.6, pp.1252–1261 (2011).
 [15] Nakano, T. and Goto, M.: VocaRefiner: An Interactive Singing Recording System with Integration of Multiple Singing Recordings, *Proc. SMC 2013*, pp.115–122 (2013).
 [16] Maaten, L. and Hinton, G.E.: Visualizing data using t-SNE, *Journal of Machine Learning Research*, pp.2579–2605 (2008).
 [17] Blei, D.M., Ng, A.Y. and Jordan, M.I.: Latent Dirichlet Allocation, *Journal of Machine Learning Research*, pp.993–1022 (2003).
 [18] Kudo, T.: MeCab: Yet Another Part-of-Speech and Morphological Analyzer, available from <http://mecab.googlecode.com/svn/trunk/mecab/doc/index.html>.
 [19] Cowan, N.: The magical number 4 in short-term memory: A reconsideration of mental storage capacity, *Journal of the Behavioral and Brain Sciences*, Vol.24, No.1, pp.87–114 (2001).
 [20] Adams, R., Ghahramani, Z. and Jordan, M.: Tree-Structured Stick Breaking Processes for Hierarchical Data, *Proc. NIPS 2010*, pp.19–27 (2010).



佐々木 将人 (学生会員)

2013年早稲田大学先進理工学部応用物理学科卒業。同年より同大学院修士課程に在学。2012年より音楽情報検索、画像処理、インタラクションに関する研究に従事。2013年情報処理学会第75回全国大会大会奨励賞受賞。

ACM 会員。



吉井 和佳 (正会員)

2008年京都大学大学院情報学研究科博士後期課程修了。同年産業技術総合研究所情報技術研究部門に入所。2014年京都大学大学院情報学研究科講師に就任。音楽情報処理、統計的音響信号処理の研究に従事。博士(情報学)。



中野 倫靖 (正会員)

2008年筑波大学大学院図書館情報メディア研究科博士後期課程修了。博士(情報学)。現在、産業技術総合研究所主任研究員。日本音響学会会員。2009年情報処理学会山下記念研究賞(音楽情報科学研究会)、2013年 Sound

and Music Computing Conference (SMC2013) The Best Paper Award 等各受賞。



後藤 真孝 (正会員)

1998年早稲田大学大学院理工学研究科博士後期課程修了。博士(工学)。現在、産業技術総合研究所情報技術研究部門首席研究員兼メディアインタラクション研究グループ長。IPA 未踏 IT 人材発掘・育成事業プロジェクト

マネージャー、情報処理学会理事等を兼任。日本学士院学術奨励賞、日本学術振興会賞、ドコモ・モバイル・サイエンス賞基礎科学部門優秀賞、科学技術分野の文部科学大臣表彰若手科学者賞、情報処理学会会長尾真記念特別賞、星雲賞等、42件受賞。



森島 繁生 (正会員)

1987年東京大学大学院工学系研究科博士課程修了。工学博士。同年成蹊大学工学部専任講師。1988年同助教授。2001年同電子工学科教授。2004年から早稲田大学先進理工学部応用物理学科教授、現在に至る。1994年から

1995年トロント大学コンピュータサイエンス学部客員教授、1999年から2014年明治大学非常勤講師、1999年より2010年国際電気通信基礎技術研究所客員研究員。2010年より2014年NICT招聘研究員、現在、新潟大学非常勤講師、早稲田大学デジタルエンタテインメント研究所所長を併任。1991年電子情報通信学会業績賞、2010年電気通信財団テレコムシステム技術賞受賞。画像電子学会副会長、日本顔学会理事、芸術科学会理事、画像2016年画像電子学会大会長、SIGGRAPH ASIA 2015 Workshop/Partner Event Chair。日本音響学会、映像情報メディア学会、日本心理学会、IEEE、ACM 各会員。