

# 利用者の印象に基づく音楽レコメンドサービス —適切な印象語を得るためのユーザインタフェースの考察—

韓語佳<sup>1</sup> 中野美由紀<sup>2</sup> 小口正人<sup>1</sup>

**概要:** 近年, 音楽配信サービスが普及し, インターネット上には数百万曲以上の楽曲が配信され, また新曲も多数発表されている. 一方で, 音楽は単なる趣味に留まらず, 音楽療法からショッピングセンターの背景まで生活の中で様々な利用されている. 多数の楽曲を利用されるシーンに併せて適切に選ぶことは難しい. そこで, 音楽の利用に合せた音楽推薦手法が期待されている. 本研究では, 新しい技術を使いにくい後期高齢者を対象とした音楽推薦手法として, 感情空間上にマッピングされた音楽DBを用い, 対話型ロボット, エージェントなどを通じて得られるその場の雰囲気やユーザの感情に適切な音楽推薦の新たな方式について検討し, 後期高齢者から適切な印象語(感情語)を得られる仕組みについて考察する.

## Music recommendation service based on user impressions - Study of user interface acquiring appropriate impression words -

YU-JIA HAN<sup>1</sup> MIYUKI NAKANO<sup>2</sup> MASATO OGUCHI<sup>1</sup>

### 1. はじめに

近年, 音楽配信サービスが普及し, インターネット上には数百万曲以上の楽曲が配信され, また新曲も多数発表されている. 一方で, 音楽は単なる趣味に留まらず, 音楽療法からショッピングセンターの背景まで生活の中で様々な利用されている. 多数の楽曲を利用されるシーンに併せて適切に選ぶことは難しい. そこで, 音楽の利用に合せた音楽推薦手法が期待されている.

また, 一人暮らしの後期高齢者や幼い子供などインターネットの利用は難しいユーザを対象に, 介護ロボット, スマートスピーカーなどを用いた簡易な対話サービスを通じ, その時々ユーザの気分(印象)に合せ, ゲーム, 動画や簡単な運動推薦等のサービスが検討され始めている.

すでに, 我々は感情空間上にマッピングされた音楽DBと印象語の関係について, アンケート調査による印象語と楽曲の位置が適切であることを確認している[5]. この結果を利用し, 現在, 我々は特定から利用したいシーン, 気分に合わせて音楽推薦を行うシステムの構築を目指している. 特に, 新しい技術を使いにくい後期高齢者を対象とした音楽推薦については, 今後の高齢化社会において重要なサービスとして期待されている. 一方で, 後期高齢者や幼児な

ど電子機器を使いこなしていないユーザに対しては, ユーザインタフェースの工夫が必要となる. スマートスピーカや対話型介護ロボットが家庭内で普及し始めている現在, 対話を楽しみながらその場の雰囲気に合せた, あるいは, 気分に合わせて印象語を得る環境が整ってきている. そこで, 対話型介護ロボットなどにおける短い対話の中から, 適切な印象語を抽出し, それを利用することを検討したい. まずは, 後期高齢者等のユーザに特有な音楽推薦シーンを検討し, 既存の対話型エージェントを基に, 感情空間上にマッピングされた音楽DBを検索するための印象語の獲得および利用時の範囲等を検討する. さらに, 少ない対話でえられる印象語に加え, 得られた印象語をより詳細な範囲に限定するための音楽推薦に特有の情報と何か等について考察を行う.

本稿は, 2節にて従来の感情空間にマッピングされた音楽DBおよび印象語, 感情空間上の位置の適切さについて, 従来の研究成果をまとめるとともに, 介護ロボットの現状について述べる. 3にて, 後期高齢者に向けた音楽推薦サービス「元気フクロウ」の概要について紹介し, 4節「元気フクロウ」のDB概要で音楽推薦に利用するデータの管理について紹介し, 5節にて音楽推薦の流れとデータの関係について詳述し, 6節で本稿のまとめを述べる.

### 2. 感情空間と音楽データベース

RussellらはArousalとValenceの2次元によって感情を表すArousal-Valence(AV)空間を提案した. ArousalとValence

<sup>1</sup> お茶の水女子大学

Ochanomizu University

<sup>2</sup> 津田塾大学

Tsuda University

はそれぞれ[-1, 1]の実数を取る。Arousal は-1 に近づくにつれて落ち着いた(calm)感情を表し、+1 に近づくにつれて興奮した(excite)感情を表す。Valence も同様に-1 に近づくにつれて、ネガティブな感情を表し、+1 に近づくにつれてポジティブな感情を表す。

図 1 では、印象語が AV 空間上にプロットされている例を示す。AV 空間上にプロットされる各印象語の座標はユーザスタディの結果によるものであり、原点からの角度が感情の意味合い、長さが感情の強さを意味する。

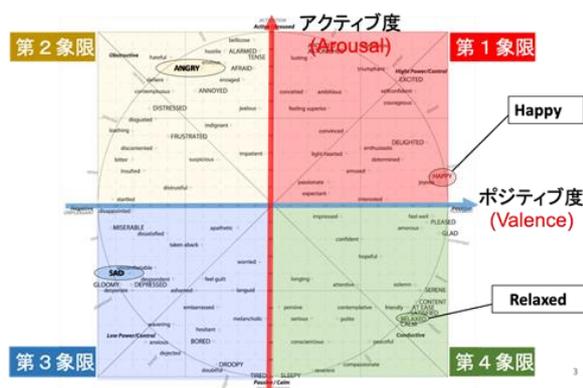


図 1 AV 空間

## 2.1 音楽データベース

音楽と感情空間の位置などについては、従来から楽曲の内容を利用する検討がなされてきている。しかし、楽曲数には限界があり、多くて数百程度の曲による検討が行われているだけである。しかしながら、音楽配信サービスなどでは 100 万曲におよぶ楽曲が提供されている。そこで、Spotify が提供する音楽の定量的評価値 (Arousal, Valence 等) を利用し、印象語 (楽しい, 悲しい等) 等の感情空間と音楽の関係について詳細に検討を行っている。[3, 5]

Spotify は数十万曲の楽曲を提供しているが、全ての音楽を調べることに限界がある。そこで、図 2 に示す通り、研究[5]では中国語及び日本語の曲を十曲選び、Spotify が提供する音楽データの数値により感情空間上にマッピングを行った。

中国語の曲については、合計 37 人の中国人 (28 人) と日本人 (9 人) からアンケートの回答が得られた。

日本語の曲については、合計 35 人の中国人 (23 人) と日本人 (12 人) からアンケートの回答が得られた。

この結果を用い、該当する楽曲と感情空間に対応する印象語と位置の適切さについて尋ねた。

表 1 と表 2 は選択した曲と AV 空間に対応する印象語である。

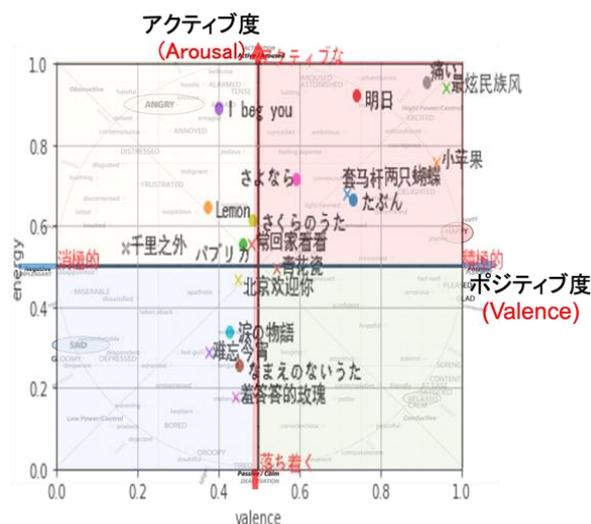


図 2 感情空間上の中国と日本曲の位置

表 1 選択した中国曲と対応する印象語

曲名	印象語
● 套马杆/Horse Pole - Namka Tso	light hearted, convinced, enthusiastic
● 两只蝴蝶/Two Butterflies - Pang Long	light hearted, convinced, enthusiastic
● 小苹果/Little Apple - Chopstick Brothers	elated
● 最炫民族风/The Most Dazzling Folk Style - Luo Yan Si	high power/control, excited
● 常回家看看/Often Go Home to Have a Look - Gong Yue	a little active
● 难忘今宵/Can't forget tonight - Gong Yue	feel guilt
● 青花瓷/Blue and White Porcelain - Jay Chou	a little positive, impressed
● 羞答答的玫瑰静悄悄地开/A Shy Rose Is Silently Blooming - Hu Xia	melancholic
● 千里之外/Far Away - Jay Chou	distrustful
● 北京欢迎你/Beijing Welcomes You - Gong Yue	a little low power/control

表 2 選択した日本語曲と対応する印象語

曲名	印象語
● たぶん - YOASOBI	enthusiastic, light hearted
● Lemon - 米津玄師	impatient, suspicious
● バブリカ - 米津玄師	a little active and a little negative
● 明日はきっといい日になる - 高橋優	triumphant
● I beg you - Aimer	afraid
● なまめのないうた - DATEKEN	languid
● さよならごっこ - amazarashi	convinced
● とても痛い痛がりたい - EZFG, TOUYU	high power/control, excited
● さくらのうた - 高橋優	a little active
● 涙の物語 - 有里知花	worried

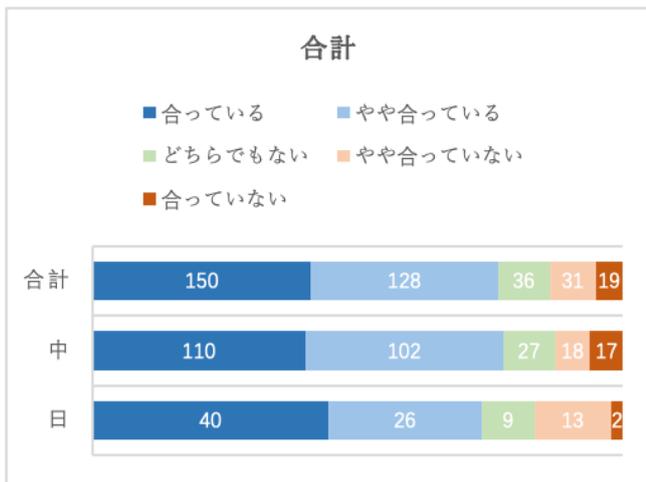


図3 中国語の十曲合計の位置の適切さ

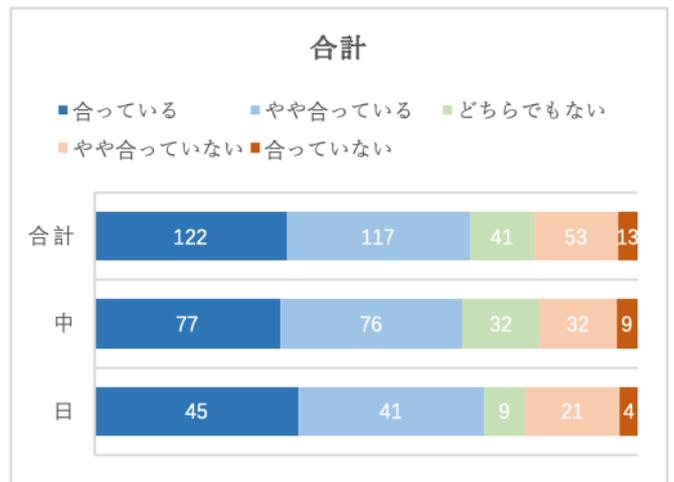


図5 日本語の十曲合計の位置の適切さ

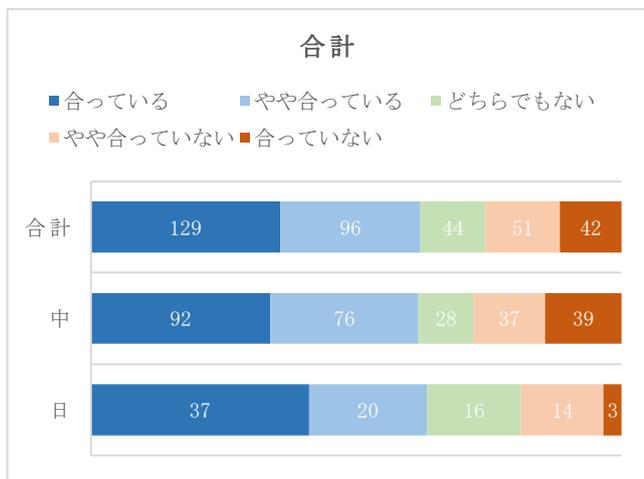


図4 中国語の十曲合計の印象語の適切さ

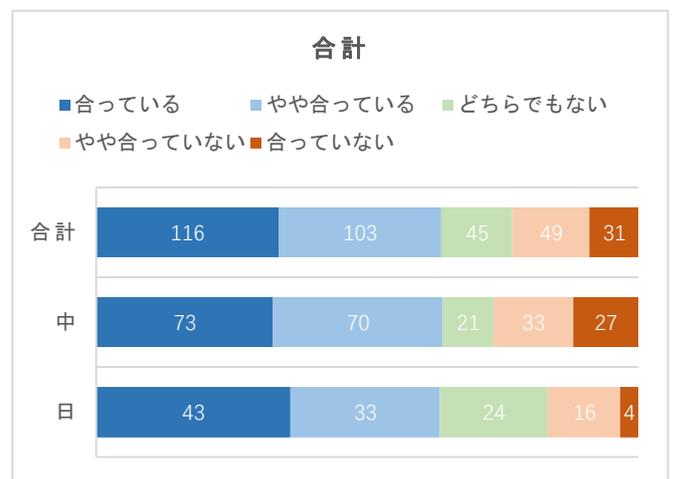


図6 日本語の十曲合計の印象語の適切さ

中国語と日本語の各 10 曲の回答を合計した結果によると、中国語の曲に関して、76.4%以上の方が、曲が感情空間上の位置が「合っている」または「やや合っている」と答えた(図3)。また、62.2%の方が、曲とそれに対応する印象語とは「合っている」または「やや合っている」と答えた(図4)。

日本語の曲に関しては、69.1%以上の方が、曲が感情空間に散布している位置が「合っている」または「やや合っている」と答えた(図5)。また、63.7%の方が、曲とそれに対応する印象語とは「合っている」または「やや合っている」と答えた(図6)。

アンケートの結果に基づき、感情空間における楽曲の位置及び印象語に関する考察を行い、Spotify が提供する Energy と Valence の値を用いた感情空間上の位置として適切であることを確認した。

## 2.2 見守り対話型ロボットと音楽配信サービス

例えばドコモ AI エージェント API [6]を利用したオートボックスセブン株式会社が発売した見守り AI ロボットブックはお年寄りと日常会話ができる。それを通じて、対話の内容や、お年寄りの安否、体調をもらえる。しかしながら、後期高齢者向けの対話型ロボット、エージェントとユーザに適した音楽推薦を合わせたサービスはまだ無い。

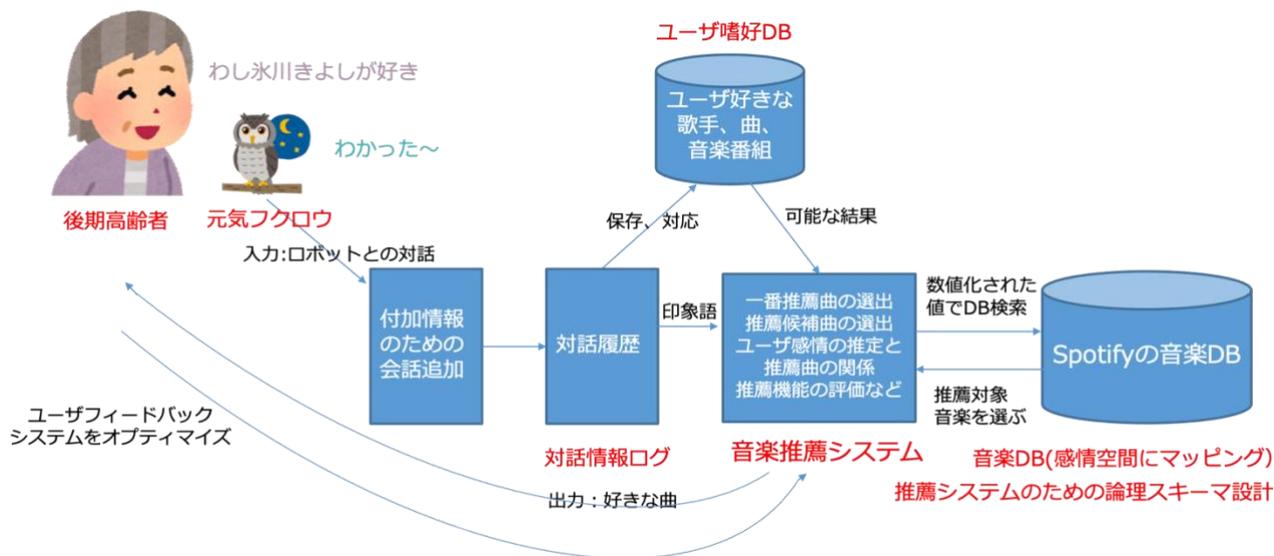


図7 音楽推薦システム「元気ふくろう」の構成図

また、音楽配信サービスそのものやスマートスピーカーでは、電子機器を使いこなせる能動的な音楽検索は可能でも、対話などの情報からその場に適切な音楽推薦を行うサービスなどは実現されていない。

### 3. 利用者の印象に基づく音楽レコメンドサービス：元気ふくろう

2節で述べたように後期高齢者向けの対話型ロボット、エージェント等のサービス、あるいは、大規模音楽配信サービスが運用されているなか、この二つを融合し、ユーザーの雰囲気や嗜好にあわせ、現在配信されている音楽の中から適切な音楽推薦を行う研究はない。我々は電子機器の操作に不慣れた後期高齢者に向けて、対話による状況把握とユーザーの嗜好に合わせた音楽推薦として感情空間にマッピングされた音楽DBを用い、より適切な音楽を提示することを目指している。

#### 3.1 サービスの概要

サービスは個人の嗜好を反映するために、対話中に現れる音楽への好みなどを収集し、気分を表す印象語に加え、個人の趣味を反映することで、適切な音楽の推薦を行う。

図7に「元気ふくろう」のシステム構成図を示す。

図にあるように、「対話エージェント機能」を持ったロボットなどを通して、後期高齢者との対話を進める。その中で、「さびしい」「楽しい」などの気分を表す印象語を得ると共に、「どんな音楽が好きですか」などの会話を投げかけることで、適切な音楽の推薦に向けたユーザーの嗜好、好きな音楽の年代などの情報収集の工夫も検討している。

「元気ふくろう」の中核となる音楽推薦システムの機能は主に、もらった対話履歴でユーザーの好みの更新と対応、

対話で得られる感情語でユーザー今の感情推定、ユーザーの情報DBで得られるユーザーの好みと対話で得られるユーザー今の気分の推定、ユーザー今の気分を利用してDBから候補曲の検索、一番推薦曲の選出、推薦候補曲の選出、推薦精度の評価である。

音楽DBは感情空間へのマッピングとして[3]で提案されたSpotifyの数値を利用すると共に、ユーザー嗜好を反映させるための特徴量もあわせて保持する。詳しくは次の4節にて説明する。また、ユーザーDBではユーザーの好きな曲のジャンル、歌手などの嗜好を保持する。

さらに、これらの推薦を通じて実際に聞かれた曲、あるいは、聞いた曲の感想などの結果をユーザーDBとして蓄積する。

元気ふくろうは2.2で紹介した対話型ロボットやスマートスピーカーサービスなどを利用して、実際の音楽データと合わせて、お年寄りが簡単に好きな音楽を楽しめ、より幸せな生活を送れることを目指している。

#### 3.2 対話から得られる印象語

一人家にいるお年寄りを対象に、ロボットとの会話のなかから印象語を取り出す。例えば、「寂しいね。娘と話したい」から「寂しい」という印象語を取る。「さっきさしぶりの友たちと話した。楽しかった！」の話から、「楽しい」との印象語を取り出す。似たようなシーンの対話から、印象語を取り出す。[7]

#### 3.3 音楽推薦の流れ

音楽推薦を行うための音楽DBなどの設計を行うために、以下のような利用パターンを検討した。

### 3.3.1 具体的な曲や具体的な言葉が指示されている時

「空の曲を放送して」などとの具体的な曲名、歌手、場所などが想定される場合について考える。

このような場合は、ユーザ DB を参照し、従来から好んで聞く曲などの傾向を合わせて推薦を行う。

入力としては、『空』が、大体二つのパターンに分ける。

#### ① 「空を聴きたい」

この場合、空は曲名、番組、歌手の名前の可能性が高い。

#### ② 「空の曲を聴きたい」

この場合、「空」は番組名、歌手の名前の可能性、あるいは空をイメージした曲などの可能性が考えられる。

このパターンに分けて、ユーザが望んでいる音楽として可能性の高いものの推薦をめざす。

まずユーザ情報 DB から、①で想定される曲、歌手等の関連する音楽からまず検索し、候補曲のなかから、Top result を提示する。そこで複数の候補曲がある場合には、例えば、「すりの空中分解ですか？」と、ユーザに質問する。ユーザから肯定的な回答が得られた場合には、該当する曲を流す。また、ユーザ情報 DB の嗜好情報として「空」という単語や流した曲、現在の対話から判断される雰囲気（明るい、楽しい、気分よい、寂しい等）をフィードバックする。

否定的な回答がかえってきた場合、候補曲の中から、ランクの高い、つまり、ユーザが好きという可能性が高い曲や関連する内容（番組か、歌手か）を選択し、ロボットはユーザに「すみません。歌手の空ですか？曲名は空ですか？」など、候補を絞ると質問する。

この場合、お年寄りが好きな可能性が高い条件も限定して検索することも可能である。

ユーザの嗜好として、80年代に流行した歌手や、曲と番組から検索し、人気度などの順位も利用可能である。

対話を通じて、「空」がユーザにとって具体的に何をさすのかが確認できれば、さらに詳細な検索を行う。

① アーティスト（歌手や作曲家）に当たる場合、音楽 DB からアーティストで検索する。いくつかの候補を示し、曲を選んでもらった後に順番に流す。同時に、曲の人気度とユーザ情報 DB とも合わせてユーザが聴きたい曲を推測してみる。

② 曲名に当たる場合は、①と同様に、人気度とユーザの好み DB と合わせて、それでもらった順位が一番の曲をひとつ流す。

③ その他の場合：「空」が番組を意味する場合もある。。お年寄りにはあまりアルバムを聞かない可能性もあるので、音楽 DB およびユーザ情報 DB の設計では、ユーザの嗜好として、単一の曲なのか、アルバムなのか、番組に該当するのかを検討している。

推薦した後、曲が終わった時点で、フィードバックをもらうために、ロボットがユーザに「この歌手が好きですか」

また「この曲が好きですか」と対話を再び開始する。その結果を音楽推薦の適切さの評価と共に、ユーザ情報 DB の更新に利用する。

また、曲を流している途中に、ユーザから「曲を変えて」などの指示が出たときは、今もうこの曲を聴きたくない判断し、ユーザ情報 DB にフィードバックにする。

### 3.3.2 具体的な指示がない時

ユーザの気分とその時の雰囲気の推測に基づき、システムからユーザに「曲を聞かないの」のような話をかけ、歌を聴くことを勧めることが出来る。この場合、すでに対話の中で得られた感情語を用い、その感情語に合わせた楽曲の推薦を検討する。

雰囲気にあわせた楽曲の推薦は印象語（感情語）を用いることで、感情空間上にマッピングされた音楽 DB から推薦することが可能である。本研究では、音楽推薦の手法について検討を行っているため、対話から得られる雰囲気、感情についてはすでに得られていると仮定している。そこで、対話から想定される雰囲気が、明るい感じ（幸せ、うれしい、明るい等）なら、明るい曲を推薦する。元気がない曲なら、元気が出る曲を推薦する。会話が少ない場合には、静かな曲を推薦したり、イライラしたり、怒っている感じでなら、穏やかな曲を推薦することを検討している。

この場合は、雰囲気あるいは感情しか分からないため、具体的な検索は行わず、感情空間上の音楽 DB を感情語の位置から検索することとなる。

① 入力としては、感情語、ユーザが好きな曲のジャンルと好きな曲の年代である。

② 出力としては、感情空間上の座標位置にそのまま該当する曲がない場合もあり、また、その感情語に適した曲は座標位置近くに多くあると考えられる。そこで、座標を中心とした適切な距離の中にある複数の曲を候補として抽出する。

③ 候補曲の中から、ユーザの嗜好にあった適切な曲を取り出す。このとき、曲のジャンルと発表された日も考慮する

### 3.3.3 その他の場合

ユーザが「いい曲を放送して」と指示した場合には、従来のユーザ嗜好と似た曲を、3.3.1 と 3.3.2 の流れにあわせて候補曲を検索し、感情語を利用しながら、ユーザ情報 DB の情報も合わせて推薦する。

この場合の結果は、ユーザのよく聴く音楽のジャンルや、歌手、曲の年代など、ユーザの特定の属性を利用して曲を推薦できる。

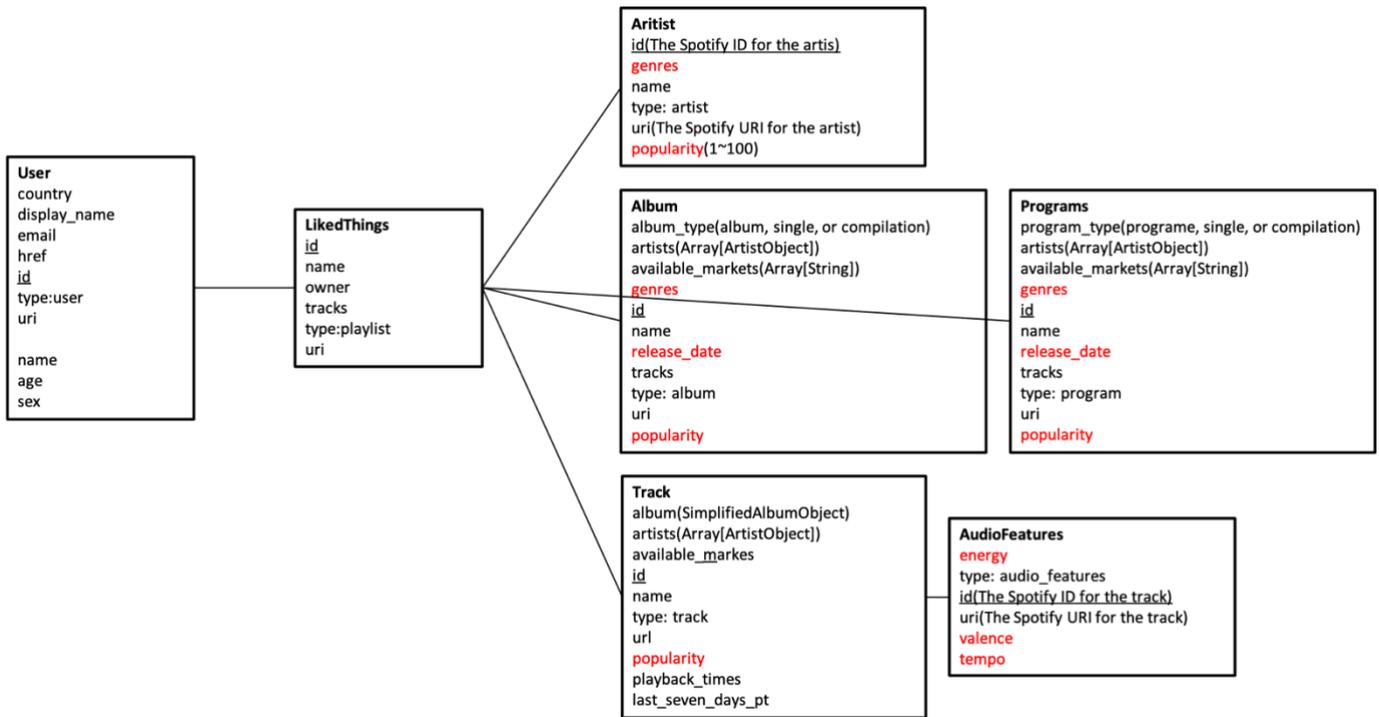


図 8 音楽推薦システムの DB 構成

## 4. 元気フクロウの DB の構成

元気フクロウの DB 構成を図 8 に示す。

### 4.1 ユーザ嗜好 DB

ユーザの嗜好などを対話から取り出すための機能が必要であるため、スキーマ設計のときに、ユーザの好みを表現することを主な目的として、前述の曲推薦事例に基づき、音楽推薦検索に必要な属性を選択した。

3.3 に音楽を推薦する三つの流れをよりスムーズに検索するために、ユーザが好きなものをデータスキーマの中心とした。その右に相応なユーザが好きな歌手、歌、テレビ番組、アルバムを右に接続する形にする。

### 4.2 音楽 DB

Spotify から得られた音楽情報の中で、ユーザの好みを表現することを主な目的として、推薦する音楽の検索に必要な属性を選択した。

それぞれは、まず曲の特徴量：valence, energy, tempo である。

valence と energy は先行研究に感情空間の座標を特定する時に使った。

tempo は今回 valence と一緒に、曲の感情がプラスかマイナスかを区別するのに使う予定である。

これらの曲の特徴量は Spotify のデータの中で、曲ごとに AudioFeatures という特徴量のデータが曲に関連するので、Track というテーブルごとに右に AudioFeatures という特徴

量に関連する。

そして、推薦曲の順番を確定する時に Spotify の人気度のデータも使うので、アルバムとアーティストと曲の人気度を表す popularity を選択した。

その後、特定の人たちに好きな曲を推薦したいため、曲のジャンルを表す genres を選択し、特定の人たちが好きな音楽のジャンルで曲を特定できる。

最後に、release\_date も使って、特定の人たちが好きな音楽の特定の年代で曲を特定できる。

一方では、ユーザの好みに合わせる音楽推薦システムの機能を実現するには、他のデータも記録する必要があるので、playback\_time というデータを追加して歌が聞かれる総回数を記録する。同じように、last\_seven\_days\_pt というデータを追加して歌が過去七日間に聞かれる総回数を記録する。

また、お年寄りが曲を聞いているときは、アルバム別で曲を聴くことより、音楽番組ごとに曲を聞くことが多いと考えられるので、アルバムと同じように、Programs というデータテーブルも追加した。

## 5. 音楽 DB を用いた推薦処理

音楽推薦するシステム元気フクロウが曲を推薦することによく使われている Spotify A P I からもらったデータは五つある：valence, energy, genres, release\_date, popularity, tempo. (ジャンル、発表された日、人気度、テンポ) 表 3 の中に赤い字で表記している。

音楽DBを用いた推薦処理の流れは、3にも紹介した通り、図7音楽推薦システム「元気ふくろう」の構成図の流れと同じように、ロボットとの対話を入力し、会話の内容でユーザが具体的な欲しいものを判断できない場合、付加情報のために会話を追加する。

そして、対話履歴を対話情報ログに記録して、そこ内容から感情語を取り出して音楽システムに提出し、ユーザ情報DBのユーザが好みの歌手、曲、音楽番組の情報と対応して、その後可能な結果を音楽システムに提出し、ユーザの好みの情報も更新する。

音楽システムは印象語とユーザ情報DBからユーザが好きな可能性があるものをもって、ユーザが欲しい曲の特徴量を推定し、それを利用してユーザ好みDBにない曲をSpotifyの音楽DBの中で曲を検索する。

そして、SpotifyDBは推薦対象になる曲を音楽推薦システムに提出しする。

その後、音楽推薦システムは全ての推薦曲の中から、一番推薦する曲を選出し、推薦候補曲も複数選出し、ユーザに曲を提出する。

最後に、ユーザは推薦してもらった曲が好きかないかを音楽推薦システムにフィードバックして、音楽推薦システムを最適化する。

それぞれのデータについて、もらった感情語と雰囲気をもとに valence, energy に変換して、AV空間に射影し、その近いところに対応する曲を探せる。

genres を使ってユーザーが好きな曲のジャンルを中心に検索できる。

release\_date を使って、特定のユーザーがよく好きな年代の曲を特定して、もっと適切な曲を検索できる。

tempo を使って、valence を補助して曲の感情はプラスかマイナスかに利用する。

## 6. まとめ

感情空間上にマッピングされた音楽DBを用いることで、電子機器などに不慣れな後期高齢者に向けた音楽推薦サービス「元気フクロウ」の提案を行った。印象語に合った音楽推薦サービスの研究は行われているが、従来の研究では多数の楽曲を用いた事例はない。

本稿では多数の音楽を配信している Spotify のデータを基に、感情空間上にマッピングした音楽DBを構築し、その上で後期高齢者に向けた音楽推薦サービスシステムの構築を目指した。

多種多様な音楽の中からユーザの嗜好に合わせた音楽、そのときの感情に合わせた音楽を推薦するために、音楽DBの構成を検討するとともに、どのようなインタフェースを用いれば、より適切な推薦音楽候補を選べるかについて、その手法を検討した。

今後は、システムの実装、DBの構築を進めるとともに、

気分（印象語）に関わる入力方法を容易かつ自然に行う手法について検討したい。

「楽しい音楽を聞きたい」、「誰々の曲を聞きたい」などの簡単な会話をもとに、お年寄りが容易に好きな音楽を視聴できることをめざしたい。

## 参考文献

- [1] Russell, J. A. (1980), "A circumplex model of affect. Journal of Personality and Social Psychology, 39(6), 1161-1178.
- [2] PALTOGLOU, Georgios, THELWALL, Michael, "Seeing stars of valence and arousal in blog posts", IEEE Transactions on Affective Computing, 2013, 4.1: 116-123.
- [3] 中野他「利用者の印象に基づく音楽レコメンドサービス：画像を利用した利用者指向の音楽レコメンド手法の考察」, 情報処理学会第81回全国大会, pp.383-384, 2019年2月.
- [4] 迫木他「画像の印象に合った音楽の推薦システムの提案」, DEIM2018-P2-3
- [5] 韓他「利用者の印象に基づく音楽レコメンドサービス：音楽データベースの感情空間へのマッピングに関する考察」, DEIM2021-D25-3
- [6] 大庭他「ドコモ AI エージェント・オープンパートナーイニシアティブ：多目的対話プラットフォームの開発」, IPSJ2017-SLP-118(10)
- [7] 大竹他「評価表現による印象推定と傾聴型対話システムへの応用」, 知能と情報（日本知能情報ファジィ学会誌）, Vol.26No.2, pp.617-626, 2014年