

アニメの主題歌による類似アニメ検索の検討

本間 直人^{†1,a)} 植村 あい子^{†1,b)} 北原 鉄朗^{†1,c)}

概要：本論文では、アニメ主題歌を用いて類似アニメを検索する手法を提案する。通常、アニメには主題歌が設定されているが、アニメの主題歌にはアニメ作品の雰囲気やテーマに沿った曲調の楽曲が採用されている傾向がある。そのため、アニメ主題歌間の類似度を測ることで、アニメ間の類似度も測ることができると考えられる。本研究では、類似楽曲検索法を用いてアニメ主題歌間の類似度を測り、同時にアニメ作品の雰囲気やテーマを表すと考えられるジャンルの一致を調査する。アニメ主題歌間の類似度を測る際に主題歌から得られる情報として、音響特徴量であるメル周波数ケプストラム係数や歌詞を特徴量とすることで、6割程度のジャンルの正解率を得ることができた。一方で、アニメ作品にはその内容を簡易的な文章で表したあらすじが設けられている。あらすじはアニメの内容を文章で直接表現しているため、類似アニメを検索する際に有用な特徴量であると考えられる。そこで、word2vecによる文章間類似度算出法を用いてあらすじ間の類似度を測った結果、8割程度のジャンルの正解率であった。

1. はじめに

アニメは今や日本を代表する文化産業の1つである。しかしアニメ作品に対するコンテンツベースの検索や推薦の研究は始まったばかりである [1]。通常、アニメには主題歌が設定されているが、アニメの主題歌にはアニメ作品の雰囲気やテーマに沿った曲調の楽曲が採用されている傾向がある。そのため、主題歌に対する類似楽曲検索 [2][3] や word2vec [4] による歌詞に対する文章間類似度算出法を用いることで、間接的な類似アニメ検索を実現できると考えられる。本稿では、文献 [3] を基にした文献 [5] の類似楽曲検索手法、word2vec [4] による歌詞に対する文章間類似度算出法を用いて類似アニメ検索を行う。一方で、アニメにはアニメ作品の内容を簡易的な文章で表したあらすじが設けられている。アニメ作品のあらすじに対する文章のベクトル化を行い、文書間類似度算出法を用いたあらすじによる類似アニメ検索を行う。そして、アニメ主題歌、歌詞を用いた類似アニメ検索とあらすじを用いた類似アニメ検索3つの手法の精度を検証する。

2. アニメ主題歌に対する類似楽曲検索手法

文献 [5] の手法を用いてアニメ主題歌に対する類似楽曲検索を行う。あるアニメ主題歌が入力されると、特徴抽出を行う。別途用意したアニメ主題歌のデータベース (DB)

の各曲に対しても同様の特徴抽出を行い、各曲と入力楽曲との距離を求め、距離が小さい順に n 曲出力する。

2.1 前処理

入力楽曲および DB 中の各曲の音響信号 (MP3) をサンプリング周波数 16kHz にダウンサンプリングした後、音響信号の中心から 60 秒分を抽出する。

2.2 メル周波数ケプストラム係数 (MFCC) の抽出

10ms 毎に 20 次元の MFCC ベクトルを抽出する。音響信号が 60 秒のため、6000 個のベクトルが得られる。

2.3 ベクトル量子化によるシグネチャの計算

6000 個のベクトルを k-means アルゴリズムを用いて 16 個のクラスタにまとめる。各クラスタを正規分布と仮定して平均ベクトルと分散共分散行列を求め、それを新たな特徴ベクトル (シグネチャ) とする。よって、楽曲 P の特徴量は以下のような式で表される。

$$P = (\mu_{p1}, \Sigma_{p1}, \omega_{p1}), \dots, (\mu_{pm}, \Sigma_{pm}, \omega_{pm}) \quad (1)$$

ここで、 μ_{pi} は平均ベクトル、 Σ_{pi} は分散共分散行列、 ω_{pi} は重み、 m はクラスタの数である。

2.4 シグネチャ間の距離の計算

2.3 節で得られたシグネチャは正規分布と仮定されているため、分布間の距離を測ることができる Earth Mover's Distance (EMD) を用いて入力楽曲と各楽曲とのシグネ

^{†1} 現在、日本大学

a) homma@kthrlab.jp

b) uemura@chs.nihon-u.ac.jp

c) kitahara@chs.nihon-u.ac.jp

チャ間の距離を計算する。

2.5 結果の出力

EMD が小さい順に n 曲 (現在の実装で 5 作品) 出力する。

3. あらすじ, 歌詞に対する文章間類似度算出法

word2vec を用いてアニメのあらすじ, 歌詞に対する文章間類似度算出を行う。あるアニメのあらすじ, 歌詞が入力されると, あらすじ, 歌詞内の各単語のベクトル化表現を行い, 各単語ベクトルの平均値を求め, 1 つの文章ベクトルとする。別途用意したアニメのあらすじ, 歌詞のデータベース (DB) の各あらすじ, 歌詞に対しても同様の単語のベクトル化表現を行い, 1 つの文章ベクトルとする。各あらすじ, 歌詞と入力したあらすじ, 歌詞とのコサイン類似度を求め, 類似度が高い順に n 作品分のあらすじ, 歌詞を出力する。

3.1 前処理

入力するあらすじ, 歌詞および DB 中の各あらすじ, 歌詞に対して, NEologed[6] を辞書に設定した MeCab[7] で形態素解析を行い, 分かち書きをする。

3.2 word2vec のパラメータの設定と学習

本研究では, 単語ベクトルの次元数を 200, 単語の最小出現回数を 1, ウィンドウサイズを 15 と設定し学習を行った。

3.3 あらすじ, 歌詞の文章ベクトル化

あらすじ, 歌詞を構成する単語 (名詞, 動詞, 形容詞に限定) のベクトルの平均を当該あらすじ, 歌詞のベクトルと見なす。

3.4 あらすじ, 歌詞間のコサイン類似度の算出

3.3 節で得られたあらすじ, 歌詞の文章ベクトルを用いて入力したあらすじ, 歌詞と DB 中の各あらすじ, 歌詞とのコサイン類似度を計算する。このコサイン類似度を求める計算式は以下のように表される。

$$\cos(\vec{q}, \vec{d}) = \vec{q}\vec{d} = \sum_{i=1}^V q_i d_i \quad (2)$$

ここで, \vec{q} は入力したあらすじ, 歌詞, \vec{d} は各あらすじ, 歌詞, V はあらすじ, 歌詞のベクトルの次元数である。

3.5 結果の出力

コサイン類似度が高い順に n 作品分のあらすじ, 歌詞 (現在の実装で 5 作品) を出力する。

4. 実験

4.1 実験方法

アニメ動画配信サイト「d アニメストア」[8] における表

表 1 「d アニメストア」における 11 種類のジャンル

SF/ファンタジー	アクション/バトル	ロボット/メカ
コメディ/ギャグ	恋愛/ラブコメ	スポーツ/競技
歴史/戦記	ホラー/サスペンス/推理	戦争/ミリタリー
日常/ほのぼの	青春/ドラマ	

1 の 11 ジャンルから再生数上位 20 位 (重複を除く) のアニメ作品を抽出し, その主題歌 (全てオープニング曲) と歌詞計 129 曲分とあらすじ計 129 作品分を用意した。129 作品の各曲 (各あらすじ, 歌詞) に対してその曲 (あらすじ, 歌詞) 自身を除いた DB を用意して類似楽曲検索 (あらすじ, 歌詞に対する文章間類似度算出法) の実験をそれぞれ行った。このとき入力楽曲 (あらすじ, 歌詞) のアニメのジャンル (複数あり) に含まれるジャンルの主題歌 (あらすじ, 歌詞) が得られたら正解とした。

4.2 実験結果

表 2 に

- アニメ主題歌に対する類似楽曲検索手法 (楽曲)
- 歌詞に対する文章間類似度算出法 (あらすじ)
- あらすじに対する文章間類似度算出法 (あらすじ)
- 全てランダムに出力する手法。 (30 回行いその平均値を計算)

それぞれの手法で実験を行った際の検索精度 (正解率) を示す。アニメ主題歌に対する類似楽曲検索手法では全体で 63%, 歌詞に対する文章間類似度算出法では全体で 66% のジャンルの正解率を得ることができた。一方で, あらすじに対する文章間類似度算出法では全体で 83% と高いジャンルの正解率を得ている。また, 全てランダムに出力した際は全体で 56% のジャンルの正解率であった。

詳細な検索結果として, アニメ作品「鋼の錬金術師」, 「おそ松さん」を対象とした際のアニメ主題歌に対する類似楽曲検索手法による検索結果を表 3, 表 4 に, あらすじに対する文章間類似度算出法の検索結果を表 5, 表 6 に, 歌詞に対する文章間類似度算出法による検索結果を表 7, 表 8 に示す。おおむねが入力した楽曲, 歌詞, あらすじのジャンルと一致していることが分かる。

5. 考察

5.1 あらすじの精度

表 2 を見てみると, あらすじに対する文章間類似度算出法では, 全体で 83%, ジャンルごとに見ても一部を除き 80% 以上と高い精度を出すことができた。しかし [日常/ほのぼの], [コメディ/ギャグ] いずれかのジャンルを含むものはその他のジャンルより精度が低くなっている。5.1 節では, これらの精度がなぜ低いのかを考察する。

まず [日常/ほのぼの] の検索結果を見てみると各ジャンルの総ジャンル数に対して, [スポーツ/競技][恋愛/ラブコ

表 2 ジャンルごとの検索精度 (正解率)

ジャンル	ランダム	MFCC	歌詞	あらすじ
SF/ファンタジー	68%	71%	81%	97%
アクション/バトル	63%	66%	73%	90%
ロボット/メカ	62%	68%	74%	99%
コメディ/ギャグ	49%	71%	55%	68%
恋愛/ラブコメ	57%	70%	73%	83%
スポーツ/競技	48%	43%	52%	84%
歴史/戦記	52%	46%	70%	76%
ホラー/サスペンス/推理	63%	65%	78%	88%
戦争/ミリタリー	50%	55%	55%	87%
日常/ほのぼの	32%	60%	51%	50%
ドラマ/青春	67%	70%	74%	94%
全体の平均	56%	63%	66%	83%

表 3 「メリッサ」([SF/ファンタジー][アクション/バトル][ドラマ/青春]) を入力楽曲としたときの検索結果 (下線: 入力楽曲とのジャンルの一致)

曲名	歌手名	ジャンル
YOUTHFUL	99RadioService	スポーツ/競技, 青春/ドラマ
さらば	キンモクセイ	日常/ほのぼの, コメディ/ギャグ
トライアングラー	坂本真綾	SF/ファンタジー, ロボット/メカ 青春/ドラマ
GLORY DAYS	D DATE	歴史/戦記, 青春/ドラマ
ボクにインビテーション	ジェイビー	日常/ほのぼの

表 4 「はなまるぴっぴはよいこだけ」([コメディ/ギャグ][ドラマ/青春]) を入力楽曲としたときの検索結果 (下線: 入力楽曲とのジャンルの一致)

曲名	歌手名	ジャンル
青春サツバツ論	3年E組うた担	アクション/バトル, ドラマ/青春
ノーボーイ!	Petit Rabbit's	日常/ほのぼの, コメディ/ギャグ
Fantastic dreamer	Machico	SF/ファンタジー, コメディ/ギャグ
トライアングラー	坂本真綾	SF/ファンタジー, ロボット/メカ, ドラマ/青春
Staple stable	戦場ヶ原ひたぎ (斎藤千和)	SF/ファンタジー, 恋愛/ラブコメ ホラー/サスペンス/推理

表 5 「鋼の錬金術師」([SF/ファンタジー][アクション/バトル][ドラマ/青春]) のあらすじを入力したときの検索結果 (下線: 入力楽曲とのジャンルの一致)

作品名	曲名	ジャンル
とある魔術の禁書目録 2	No Buts!	SF/ファンタジー アクション/バトル
鋼の錬金術師 FULLMETAL ALCHEMIST	again	SF/ファンタジー, アクション/バトル ドラマ/青春
クロスアンジュ 天使と竜の輪舞	禁断のレジスタンス	SF/ファンタジー, ロボット/メカ ドラマ/青春
夏目友人帳	一斉の声	SF/ファンタジー, ホラー/サスペンス/推理 ドラマ/青春
ブラックブレット	blackbullet	SF/ファンタジー, アクション/バトル 恋愛/ラブコメ

メ] の順にいずれかを含んだ作品が不正解として多く挙げられていた。一方で, [ロボット/メカ] を含んだ作品の不正解数は 0 だった。

このことから [日常/ほのぼの] のジャンルを含んでいる作品は, [スポーツ/競技][恋愛/ラブコメ] それぞれのジャン

表 6 「おそ松さん」([コメディ/ギャグ][ドラマ/青春]) のあらすじを入力したときの検索結果 (下線: 入力楽曲とのジャンルの一致)

作品名	曲名	ジャンル
銀魂 2 年目	銀色の空	コメディ/ギャグ, アクション/バトル ドラマ/青春
銀魂 1 年目	Pray	コメディ/ギャグ, アクション/バトル ドラマ/青春
銀魂 4 年目	Stairway Generation	コメディ/ギャグ, アクション/バトル ドラマ/青春
銀魂 3 年目	曇天	コメディ/ギャグ, アクション/バトル ドラマ/青春
今日から魔王	again	コメディ/ギャグ, アクション/バトル SF/ファンタジー

表 7 「鋼の錬金術師」([SF/ファンタジー][アクション/バトル][ドラマ/青春]) の歌詞を入力したときの検索結果 (下線: 入力楽曲とのジャンルの一致)

作品名	曲名	ジャンル
四月は君の嘘	光るなら	恋愛/ラブコメ, ドラマ/青春 コメディ/ギャグ
銀魂 2 年目	銀色の空	ドラマ/青春, アクション/バトル
ソードアート・オンライン 2	IGNITE	SF/ファンタジー, アクション/バトル, ドラマ/青春
アルドノアゼロ	heavenly blue	SF/ファンタジー, ロボット/メカ SF/ファンタジー,
化物語	staple stable	ホラー/サスペンス/推理, 恋愛/ラブコメ

表 8 「はなまるぴっぴはよいこだけ」([コメディ/ギャグ][ドラマ/青春]) の歌詞を入力したときの検索結果 (下線: 入力楽曲とのジャンルの一致)

作品名	曲名	ジャンル
生徒会役員共*	花咲く 最強レジェンド Days	恋愛/ラブコメ, コメディ/ギャグ
青の祓魔師	CORE PRIDE	アクション/バトル ホラー/サスペンス/推理
はじめの一歩 New Challenger	HEKIREKI	スポーツ/競技, ドラマ/青春
るろうに剣心 明治剣客浪漫譚	そばかす	アクション/バトル, ドラマ/青春
ノーゲームノーライフ	This game	SF/ファンタジー スポーツ競技, 恋愛/ラブコメ

ルを含んでいる作品と混同されやすく, [ロボット/メカ] のジャンルを含んでいる作品と混同されにくいことが考えられる。

この仮説を裏付けるために [日常/ほのぼの] に混同されやすいジャンル, 混同されにくいジャンルそれぞれの特徴語を調べ, その特徴語間の類似度を word2vec を用いて計算した。文書に含まれる単語の重要度を評価する手法の 1 つである tf-idf を用いて出力した各ジャンルの特徴語上位 10 件を表 9 に示す。また, 各ジャンルの特徴語間の類似度を計算した結果を表 10 に示す。

表 10 より [日常/ほのぼの] の特徴語と, [スポーツ/競技][恋愛/ラブコメ] の特徴語間の類似度を見てみると [ロボット/メカ] との特徴語間の類似度に比べて高いことがわかる。よって [日常/ほのぼの] のジャンルを含んでいる作品は, [スポーツ/競技][恋愛/ラブコメ] それぞれのジャンルを含んでいる作品と特徴語が類似しているため混同されやすく, [ロボット/メカ] のジャンルを含んでいる作品とは特徴語が類似していないため混同されにくいと考えられる。

[コメディ/ギャグ] の検索結果を見てみると各ジャンルの総ジャンル数に対して, [スポーツ/競技][ドラマ/青春] の

表 9 [日常/ほのぼの][スポーツ/競技][恋愛/ラブコメ][ロボット/メカ] の特徴語上位 10 件

ランク	[日常/ほのぼの]	[スポーツ/競技]	[恋愛/ラブコメ]	[ロボット/メカ]
1	青葉	戦車道	IS	IS
2	料理	バスケット	クラス	オブジェクト
3	イーグル・ジャンプ	みほ	キャラ	ネウロイ
4	入社	南郷	ピアノ	ラプラス
5	新入社員	大洗女子学園	空白	箱
6	つむぐ	履修	秀麗	地球
7	クール	かるた	官吏	ロボット
8	ラビットハウス	花道	女性	G-セルフ
9	ちの	赤木	ネトゲ	OZMA
10	かわいい	麻雀	荘	鉄華団

表 10 [日常/ほのぼの][スポーツ/競技][恋愛/ラブコメ][ロボット/メカ] の特徴語における全組み合わせ間の類似度の平均

[日常/ほのぼの]-[スポーツ/競技]	0.092321
[日常/ほのぼの]-[恋愛/ラブコメ]	0.076678
[日常/ほのぼの]-[ロボット/メカ]	0.018250

表 11 [コメディ/ギャグ][スポーツ/競技][ドラマ/青春][歴史/戦記] の特徴語上位 10 件

ランク	[コメディ/ギャグ]	[スポーツ/競技]	[ドラマ/青春]	[歴史/戦記]
1	キメる	戦車道	戦車道	陽子
2	キャラ	バスケット	バスケット	慶
3	クラス	みほ	みほ	ケイキ
4	ラビットハウス	南郷	南郷	租界
5	チノ	大洗女子学園	妖怪	秀麗
6	かわいい	履修	錬成	官吏
7	勇者	かるた	バラサイト	王
8	ネトゲ	花道	大洗女子学園	国王
9	宇宙人	赤木	履修	甲賀
10	天人	麻雀	エド	協商

表 12 [コメディ/ギャグ][スポーツ/競技][ドラマ/青春][歴史/戦記] の特徴語における全組み合わせ間の類似度の平均

[コメディ/ギャグ]-[スポーツ/競技]	0.124504
[コメディ/ギャグ]-[ドラマ/青春]	0.120666
[コメディ/ギャグ]-[歴史/戦記]	0.055548

順にいずれかを含んだ作品が不正解として多く挙げられた。一方で、[歴史/戦記] を含んだ作品の不正解数が最も少なかった。[日常/ほのぼの] と同様に [コメディ/ギャグ] に混同されやすいジャンル、混同されにくいジャンルそれぞれの特徴語上位 10 件を表 11 に、特徴語間の類似度を計算した結果を表 12 に示す。

表 12 より [コメディ/ギャグ] の特徴語と、[スポーツ/競技][ドラマ/青春] の特徴語間の類似度を見てみると [歴史/戦記] との特徴語間の類似度に比べて高いことがわかる。

5.2 MFCC の精度

表 2 を見てみると、アニメ主題歌に対する類似楽曲検索法では、全体で 63%、ジャンルごとに見ても一部を除きおおむね 60%以上の精度を出すことができた。しかし [スポーツ/競技]、[歴史/戦記] いずれかのジャンルを含むものはその他のジャンルより精度が低くなっている。5.2 節では、これらの精度がなぜ低いのかを考察する。

表 13 [スポーツ/競技][日常/ほのぼの][ホラー/サスペンス/推理] の楽曲間の EMD の平均

[スポーツ/競技]-[スポーツ/競技]	55.47138
[スポーツ/競技]-[日常/ほのぼの]	54.60162
[スポーツ/競技]-[ホラー/サスペンス/推理]	57.89729

表 14 [歴史/戦記][日常/ほのぼの][コメディ/ギャグ] の楽曲間の EMD の平均

[歴史/戦記]-[歴史/戦記]	89.62123
[歴史/戦記]-[日常/ほのぼの]	74.14599
[歴史/戦記]-[コメディ/ギャグ]	74.32233

[スポーツ/競技] の検索結果を見てみると各ジャンルの総ジャンル数に対して、[日常/ほのぼの][ホラー/サスペンス/推理] の順にいずれかを含んだ作品が不正解として多く挙げられていた。そこで [スポーツ/競技] のジャンルを含む作品の楽曲間同士、[スポーツ/競技] のジャンルを含む作品と [日常/ほのぼの] のジャンルを含む作品の楽曲間、[スポーツ/競技] のジャンルを含む作品と [ホラー/サスペンス/推理] のジャンルを含む作品の楽曲間の EMD の平均値を計算した。その結果を表 13 に示す。同様に、[歴史/戦記] の検索結果を見てみると各ジャンルの総ジャンル数に対して、[日常/ほのぼの][コメディ/ギャグ] の順にいずれかを含んだ作品が不正解として多く挙げられていた。[歴史/戦記] のジャンルを含む作品の楽曲間同士、[歴史/戦記] のジャンルを含む作品と [日常/ほのぼの] のジャンルを含む作品の楽曲間、[歴史/戦記] のジャンルを含む作品と [コメディ/ギャグ] のジャンルを含む作品の楽曲間の EMD の平均値を計算した結果を表 14 に示す。

表 13 より [スポーツ/競技] のジャンルを含む作品の楽曲間同士の EMD の平均値が、[日常/ほのぼの][ホラー/サスペンス/推理] それぞれのジャンルを含む楽曲間との EMD の平均値に比べて高い、もしくは近い値であることがわかる。

また、表 14 でも [歴史/戦記] のジャンルを含む作品の楽曲間同士の EMD の平均値が、[日常/ほのぼの][コメディ/ギャグ] それぞれのジャンルを含む楽曲間との EMD の平均値に比べて高い値であることがわかる。

このことからアニメ主題歌に対する類似楽曲検索法では、正解であるはずの自身のジャンルを含む作品の楽曲より他のジャンルを含む作品の楽曲の方が出力されやすいことが [スポーツ/競技][歴史/戦記] の精度が低い原因だと考えられる。

5.3 歌詞の精度

表 2 を見てみると、歌詞に対する文章間類似度算出法では、全体で 66%、ジャンルごとに見ても一部を除きおおむね 70%以上の精度を出すことができた。しかし [日常/ほのぼの][スポーツ/競技] いずれかのジャンルを含むものはその他のジャンルより精度が低くなっている。5.3 節では、[日

表 15 [日常/ほのぼの][戦争/ミリタリー][スポーツ/競技][歴史/戦記]の歌詞間の類似度

[日常/ほのぼの]-[戦争/ミリタリー]	0.212760
[日常/ほのぼの]-[スポーツ/競技]	0.246256
[日常/ほのぼの]-[歴史/戦記]	-0.108738

常/ほのぼの]に焦点を当てて、精度がなぜ低いのかを考察する。

[日常/ほのぼの]の検索結果を見てみると各ジャンルの総ジャンル数に対して、[戦争/ミリタリー][スポーツ/競技]の順にいずれかを含んだ作品が不正解として多く挙げられていた。一方で、[歴史/戦記]を含んだ作品の不正解数が最も少なかった。そこで[歴史/戦記]のジャンルを含む作品の歌詞間同士、[歴史/戦記]のジャンルを含む作品と[日常/ほのぼの]のジャンルを含む作品の歌詞間、[歴史/戦記]のジャンルを含む作品と[コメディ/ギャグ]のジャンルを含む作品の歌詞間のコサイン類似度を計算した。その結果を表 15 に示す。

表 15 より [日常/ほのぼの] の歌詞と、[戦争/ミリタリー][スポーツ/競技] の歌詞間の類似度を見てみると [歴史/戦記] との歌詞間の類似度に比べて高いことがわかる。よって [日常/ほのぼの] のジャンルを含んでいる作品の歌詞は、[戦争/ミリタリー][スポーツ/競技] それぞれのジャンルを含んでいる作品の歌詞と類似しているため混同されやすく、[歴史/戦記] のジャンルを含んでいる作品とは歌詞が類似していないため混同されにくいと考えられる。

6. おわりに

本研究では、アニメの主題歌にはアニメ作品の雰囲気やテーマに沿った曲調の楽曲が採用されている傾向があるといった特徴から、楽曲間、歌詞間の類似度を測ることで類似アニメを検索する手法を検討した。アニメジャンルをアニメ作品間の類似度の指標とし、楽曲間の類似度を測った結果、全体として MFCC では 63%、歌詞では 66% のアニメジャンルの正解率を得られることができた。このことから、本研究の楽曲を用いた類似アニメ検索手法に於いては現段階で 6 割程度の精度で類似アニメを検索することが可能であることが分かった。

一方で、アニメのあらすじによる文章間の類似度算出法では全体として 83% と高いアニメのジャンルの正解率を得ることができたことから、あらすじを用いたアニメ推薦は有効であるということが示された。しかし、全体の精度としてみると、あらすじを用いたほうが精度は良いが、表 2 ジャンルごとの検索精度を見てみると [コメディ/ギャグ]、[日常/ほのぼの] を含むものは、あらすじを用いたときよりも楽曲を用いたときの方が同等または高い検索精度を得ていることが分かる。

今後主題歌、歌詞による手法とあらすじによる手法の統

合やリズムを考慮した類似楽曲検索を視野に入れて、類似アニメ検索の精度を高めていきたい。

謝辞

本研究は、JSPS 科研費 < 17H00749 > から支援を受けた。

参考文献

- [1] H. Cho, M. L. Schmalz, S. A. Keating, and J. H. Lee: Information Needs for Anime Recommendation: Analyzing Anime Users' Online Forum Queries, ACM/IEEE Joint Conf. on Digital Libraries (JCDL), 2017.
- [2] E. Pampalk: Computational Models of Music Similarity and their Application in Music Information Retrieval, Doctoral Thesis, Vienna University of Technology, Austria, 2006.
- [3] B. Logan and A. Salomon: A Content-based Music Similarity Function, Technical report, Compaq Cambridge Research Lab, 2001.
- [4] word2vec, <https://code.google.com/archive/p/word2vec/>
- [5] 人工知能に関する断創録 類似楽曲検索を作ろう, <http://aidiary.hatenablog.com/entry/20121014/1350211413>
- [6] mecab-ipadic-neologd, <https://github.com/neologd/mecab-ipadic-neologd>
- [7] MeCab, <http://taku910.github.io/mecab/>
- [8] d アニメストア, <https://anime.dmkt-sp.jp/>