

音楽のジャンルと印象を用いた VOCALOID クリエータの検索

VOCALOID Creator Search Using Music Genre and Impression

末吉 優
Yuu Sueyoshi

筑波大学 情報学群 知識情報・図書館学類
College of Knowledge and Library Sciences, School of Informatics, University of Tsukuba

関 洋平
Yohei Seki

筑波大学 図書館情報メディア系
Faculty of Library, Information and Media Science, University of Tsukuba
yohei@slis.tsukuba.ac.jp, <http://cu.slis.tsukuba.ac.jp/~seki/>

keywords: VOCALOID, creator search, music genre, Niconico

Summary

This paper proposes the following methods to search VOCALOID creators who publish music videos in Niconico video hosting service. For VOCALOID creator search, the user can utilize three clues: VOCALOID character name, music genre, and impressions. We defined the music genre by extending generic digital music genre with considering social tags annotated on VOCALOID music videos. We also implemented SVM-based music impression estimator utilizing viewer comments being over 0.8 points in F-values. We compared the proposal with three comparison methods in 12 search tasks and clarified the effectiveness of music genres and impressions.

1. ま え が き

近年、株式会社ヤマハ*1が開発した歌声合成技術およびその応用ソフトウェアである VOCALOID*2を使用した楽曲動画が、ニコニコ動画*3や YouTube*4などの動画投稿サービスに多数アップロードされている。VOCALOID 楽曲動画はこれらの動画サイトにおける人気コンテンツとなっており、VOCALOID 楽曲動画を収録した音楽 CD が市場に出回ったり、カラオケの人気曲として話題になったりするなど、ひとつの音楽文化として確立しつつあるといえる。本研究では、ユーザの好みを反映した検索キーを指定することにより、こうした VOCALOID 楽曲動画を制作しているクリエイターを発見するシステムを提案する。

VOCALOID 楽曲動画を制作している人々の中には、アマチュアで活動する人が多く存在し、「音楽やイラスト、映像等のさまざまなコンテンツで潜在能力が高く熱意あるクリエイターが日本には多く、従来は埋もれていたのが、1次創作やN次創作等で活躍するようになった」[後藤12]といわれている。このように、動画サイトで活躍しているクリエイターは、インターネット上で固定ファンを獲

得しながらコンテンツを配信しており、YouTube においては、人気のある動画配信者が「トップチューバー」として注目され、企業と提携して商品やサービスを生み出すという現象が起こるなど、その存在は社会的にも大きくなっている。

このような背景を踏まえ、本研究では、クリエイターが VOCALOID 楽曲動画を公表する場となっているニコニコ動画において、投稿された楽曲動画に付与されたタグと視聴者コメントを利用して、VOCALOID を使用した楽曲動画を制作し、発表しているクリエイターを、ユーザの好みを反映するかたちで検索するシステムを提案する。このシステムにより、クリエイターの使用する VOCALOID 名、よく制作する楽曲動画のジャンル、楽曲動画に対する印象を用いてクリエイターを検索することで、自分の好きな音楽に、より効果的にアクセスすることが可能となる。

本研究では、楽曲動画に付与されたタグから、その楽曲動画を制作しているクリエイター名、使用されている VOCALOID 名、楽曲動画のジャンルを判別し、視聴者コメントによって楽曲動画の印象を分類することで、クリエイターごとに制作した楽曲動画の傾向を整理し、クリエイターの検索に活用することを目指す。

本研究の貢献は、以下の3点である。

- (1) クリエータ検索という概念および検索方法を提案

*1 <http://jp.yamaha.com/>
*2 <http://www.vocaloid.com/>
*3 <http://www.nicovideo.jp/>
*4 <https://www.youtube.com/>

する。これにより、ユーザが好みのクリエイターを発見することや、楽曲動画のクリエイターに、コンテンツ制作の仕事を依頼することを支援できるようになる。

- (2) VOCALOID クリエータの検索に適した音楽ジャンルを提案し、既存ジャンルとの比較実験を通して、その有効性を示す。
- (3) クリエータの検索課題を用意し、比較実験を行うことにより、手がかりとして、音楽ジャンルや楽曲動画の印象が有効である場合を明らかにする。

本論文の構成について述べる。2 章では、楽曲動画およびクリエイターの検索に関する研究について紹介する。3 章では、提案するクリエイター検索システムの概要と、データの収集について述べる。4 章では、クリエイター検索システムの実装に必要な、楽曲動画ジャンルのラベルの定義に関する調査と、楽曲動画の印象分類について説明する。5 章では、クリエイター検索におけるランキング手法について説明する。6 章では、クリエイター検索システムについて比較実験を行い、提案手法の有効性を明らかにする。最後に、7 章で、本研究のまとめを行う。

2. 関連研究

2.1 楽曲動画の検索

VOCALOID 楽曲動画のクリエイターを探すために、Twitter 上でハッシュタグを利用して、「#ボカロ P (さん) と繋がりたい」というようなメッセージを投稿することがある。こうしたハッシュタグは、クリエイターを探したいユーザと、自分をクリエイターとして見つけてほしいユーザの双方が利用して投稿している。これらのツイートを眺めてみると、自分をクリエイターとして見つけてほしいユーザは、自分の楽曲動画へのリンクを張ることで、リンク先の楽曲動画を視聴するユーザに、ユーザが作成を依頼したい楽曲動画に対するクリエイターとしての適合性の判断を委ねているケースが多い。しかし、多様なクリエイターが存在することから、意中のクリエイターを発見することは至難の業である。したがって、クリエイターが得意とする楽曲動画を判別するための手がかりについて考慮する必要がある。また、リンク先の情報として、歌唱している VOCALOID キャラクター名 (初音ミクなど) を提示しているケースが見られることから、楽曲動画のクリエイターが得意とする VOCALOID 名がクリエイター判別の手がかりとなることが示唆される。一方で、佃ら [佃 11] は、視聴者の反応に基づいた動画を検索するシステムを提案しており、動画中に登場する人物が活躍しているシーンを「活躍シーン」として定義し、時刻同期コメントを利用して動画における活躍シーンの場所を判定することで、「ある人物が動画の後半から活躍する泣ける動画」といった検索を実現している。この研究では、動画に登場する人物による検索を行うことによって、よりユーザの要求を反映したシステムを実現している。以上を参

考にして、本研究においては、クリエイターが作成した楽曲動画に登場する VOCALOID のキャラクターに焦点を当てた検索を実現する。

一方、楽曲を検索する手がかりは、音楽ジャンル [Correa 16, Sturm 14] や印象 (あるいは感情) [金礪 13, Kim 10, 熊本 02, 大野 16, Yang 12] が利用される。音楽ジャンルを分類する手がかりとしては、音響信号 [Costa 12, Fu 11] が用いられることが多い。また、最近のソーシャルメディアの発展に伴い、ソーシャルメディア上で付与されるタグを利用して音楽ジャンルを分類する研究も提案されている [Chen 09, Wang 09]。音楽ジャンルは、Web 上の楽曲検索において有力な指標となる [土橋 06]。

Chen らは、音楽についてのコミュニティ Web サイトである Last.fm の API を利用して収集したタグを手がかりとすることにより、音楽ジャンルを分類する手法を提案した [Chen 09]。本研究で対象とする、VOCALOID 楽曲動画については、クリエイター検索に適した独自の音楽ジャンルを設定する必要があると考える。本研究では、楽曲動画の視聴者により付与されたタグを手がかりとして、音楽ジャンルを判別する方法を提案する。

楽曲動画の印象分類の手がかりとしては、音響に加えて歌詞が用いられることが多い [Hu 10, Jamdar 15, Ogino 15]。また、視聴者の反応を利用する手法としては、ソーシャルメディア上のタグ [Hu 10, Jamdar 15] を利用する手法がある。印象の分類については、Russell [Russell 80] により提案された Valence-Arousal の空間表現が広く利用されている [Soleymani 13]。

VOCALOID 楽曲動画に対する印象については、視聴者のコメントを利用する先行研究 [山本 13b] がある。山本ら [山本 13b] は、ニコニコ動画における視聴者コメントを分析し、楽曲動画を「可愛らしい」、「切ない」、「元気が出る」などの 7 種類の印象に分類する手法を提案している。素性の抽出方法としては、コメント中の形容詞および形容動詞の抽出、繰り返し文字の正規化、サビ区間のコメントの利用という 3 つの方法をとっている。また、この研究をもとにして作成された楽曲動画の印象データセット [山本 13a] が公開されている。このデータセットを利用して楽曲動画の印象を推定することにより、クリエイターがよく作成する楽曲動画が与える印象を手がかりとした検索に応用することが検討できる。

本研究では、楽曲動画の印象を推定するために、自然言語処理や情報検索の分野で使用されることの多い相互情報量 (expected mutual information) [Manning 08] を利用して素性を選択することにより、高い推定精度を実現できることを示す。また、クリエイターが制作した楽曲動画に付与されたタグや視聴者のコメントを解析することにより、VOCALOID 名に加えて、音楽ジャンルや楽曲動画の印象といった手がかりを獲得し、VOCALOID 楽曲動画のクリエイター検索における有効性について検証する。

2.2 クリエータの検索

Web 上で作品を制作・公開しているクリエイターを検索する研究として、Seki ら [Seki 13] は、イラスト共有 SNS の Pixiv*5において、イラスト群に付与されたタグから投稿者のモチーフと印象を推定し、イラスト投稿者を検索する手法を提案している。この研究においては、個々の作品に対するモチーフと印象の集合を、投稿者がどのような印象の作品を作っているかを推定することに利用している。本研究では、この研究を踏まえて、個々の楽曲動画に付与されたタグやコメント機能を利用し、クリエイターが制作する楽曲動画のジャンルや印象推定を行い、クリエイターの検索に役立つ手法を提案する。

3. クリエータ検索システムの概要

本研究で実現するクリエイターの検索では、クリエイターが制作している楽曲動画の歌唱 VOCALOID 名、音楽のジャンル、楽曲動画の印象を用いて、以下に示す手順でクリエイターを検索する。

- (1) VOCALOID 名、音楽のジャンル、楽曲動画の印象を入力する
- (2) クリエータの名前、使用している VOCALOID 名、制作する音楽ジャンル、楽曲動画の印象、投稿している楽曲動画の ID 情報を格納した、クリエイターのデータベースに問い合わせる
- (3) 入力内容とクリエイターの情報を比較し、合致するクリエイターの名前と投稿している楽曲動画集合をユーザに提示する

以上の手順をまとめた全体像を、図 1 に示す。

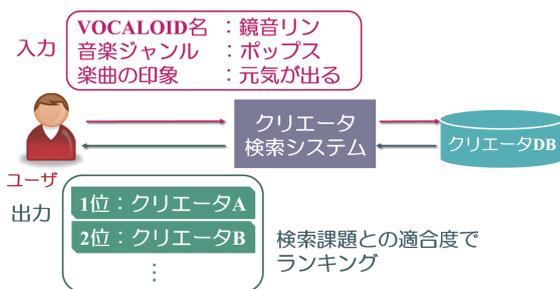


図 1 クリエータ検索システムの全体像

3.1 楽曲動画とタグデータの収集およびクリエイター名の判別

本節では、クリエイター検索システムに利用する楽曲動画ならびに楽曲動画に付与されたタグの収集方法について説明する。楽曲動画とタグは、動画提供サービスである「ニコニコ動画」*6から、ニコニコ動画 API*7を用いて、収集する。タグは、VOCALOID 楽曲動画に特有の

タグを収集することで、楽曲動画のジャンルの推定に利用する。最初に、クリエイターが制作した楽曲動画の収集方法について説明する。

- (1) まず、ニコニコ動画サービス設立時（動画 ID:sm1）から、2014 年 7 月 14 日（動画 ID:sm2400000）までの期間に投稿された動画のうち、「VOCALOID」タグが付与された 314,029 件を対象として収集する。
- (2) 次に、手順 (1) で収集した楽曲動画から、ニコニコ大百科*8の中の VOCALOID のプロデューサー一覧*9という記事に記載されている 1,561 名（2014 年 6 月 9 日時点）のクリエイター名ならびに、「オリジナル曲」という文字列をタグとして付与されている楽曲動画を選択する。その結果、1,285 名のクリエイターが作成した 22,593 件の楽曲動画が得られた。

なお、「オリジナル曲」というタグを制約として与えたのは、クリエイター名のみの抽出を行うと、「歌ってみた」や「描いてみた」などの、クリエイターが投稿したオリジナル曲に関連する他者の投稿作品も選択されてしまうという理由による。

対象外とされた 276 名のクリエイターについては、オリジナル曲を投稿していない以外に、以下の傾向が見られた。

- a クリエータ名のタグが正確に照合できていないか、表記揺れがある
- b 投稿した動画を全て削除してしまっている
- c VOCALOID 以外の音声合成ソフトを用いている

以上の手順により判別したクリエイターは、1 人あたり平均 17.6 件の楽曲動画を制作しており、最も少ないクリエイターは 1 件、最も多いクリエイターは 639 件の楽曲動画を投稿している。

クリエイターの投稿した楽曲動画数の分布を図 2 に示す。楽曲動画を 5 件以上投稿しているクリエイターは 991 名で、これは全体の約 77% であり、クリエイターの検索実験に用いるにあたって、十分な手がかりが獲得できるユーザが一定数存在することが確認できた。

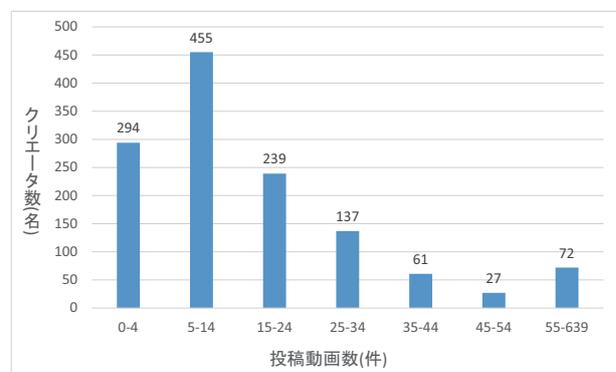


図 2 クリエータの投稿した楽曲動画数の分布

*5 <http://www.pixiv.net/>

*6 <http://www.nicovideo.jp/>

*7 http://dic.nicovideo.jp/a/ニコニコ動画_api

*8 <http://dic.nicovideo.jp/>

*9 <http://dic.nicovideo.jp/a/vocaloid> のプロデューサーの一覧

3.2 VOCALOID の名前の判別

入力として用いる、クリエイターが頻繁に使用する VOCALOID 名については、51 件のキャラクター (2014 年 8 月末までに発売されているもの) を対象として、動画に付与されているタグと照合することにより判別した。この VOCALOID 名で検索対象となる楽曲動画の約 91.6% をカバーできる。利用した VOCALOID 名のタグの一覧を表 1 に示す。

表 1 VOCALOID 名のタグの一覧

LEON	Big-AL	蒼姫ラビス
LOLA	TONIO	洛天依
MIRIAM	Lily	MAYU
MEIKO	VY1	Avanna
KAITO	ガチャポイド	ZOLA.PROJECT.KYO
SweetAnn	猫村いろは	ZOLA.PROJECT.YUU
初音ミク	歌手音ピコ	ZOLA.PROJECT.WIL
PRIMA	VY2	言和
鏡音リン	Mew	YOHIloid
鏡音レン	SeeU	マクネナナ
がくっぽいど, 神威がくぼ	兔眠りおん	MAIKA
巡音ルカ	Oliver	kokone
GUMI, Megpoid, メグポイド	結月ゆかり	杏音鳥音
SONIKA	CUL	メルリ
氷山キヨテル	BRUNO	v.flower
歌愛ユキ	CLARA	東北ずん子
開発コード miki	IA.-ARIA.ON.THE.PLANETES-	ギャラ子, galaco

3.3 音楽ジャンルの体系

VOCALOID 楽曲動画には、楽曲動画のジャンルを判別するジャンルタグが存在しており、視聴者はそのタグを頼りに、好きなジャンルの楽曲動画を探することができる。しかし、ジャンルタグの数は膨大であり、VOCALOID 楽曲動画特有のタグであるために、馴染みのないユーザにはわかりづらい表現になっているものが多い。そこで、類似するジャンルタグをより明確なジャンルラベルにまとめることで、ユーザにとって扱いやすいジャンルの分類が実現できると考え、音楽ジャンルラベルを定義し、ラベルにジャンルタグを対応付けるための調査を行う。この調査に関しては、4.1 節で詳しく述べる。

3.4 楽曲動画の印象の自動分類に利用するデータ

楽曲動画の印象に関して、山本ら [山本 13a] によって構築された、楽曲動画印象データセット^{*10}が存在する。本研究ではこのデータセットおよび楽曲動画に付与されたコメントのデータを評価データとした、LIBSVM^{*11}による機械学習により、クリエイターの制作する楽曲動画の印象を自動分類する。

学習に利用する楽曲動画のコメントデータに関しては、ニコニコ動画 API を利用して、2014 年 11 月 17 日～25 日の間に、最新の 1,000 件までのデータを楽曲動画ごとに収集した。機械学習による印象の自動分類に関しては、4.2 節で詳しく述べる。

*10 <http://nkmr.io/mood/>

*11 <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>

4. 音楽ジャンルの定義と楽曲動画の印象の自動分類

本章では、提案するクリエイター検索システムの実装に必要な音楽ジャンルの定義に関する調査と、クリエイター検索システムの実装に必要な楽曲動画の印象分類について述べる。

4.1 音楽ジャンルのラベルと対応するタグの選定

§1 目的

ユーザの好む音楽ジャンルによってクリエイターを検索するため、実験参加者である 3 名の学生により、クリエイターに付与するジャンルラベルを定義し、ジャンル分類の手がかりとなるタグを対応付ける。また、作成したジャンル分類体系を利用して、クリエイターにジャンルラベルを付与し、音楽ジャンルの自動付与について検証する。

§2 方法

(1) 実験参加者が、ジャンルラベルと対応するタグを選定

a ジャンルラベルの選定

実験参加者には、以下の指示を与えた。まず、Amazon.co.jp の「デジタルミュージック」における「ジャンル別ベストセラー」の 12 項目のジャンル (「J-POP」、「ポップス」、「ロック」、「ハードロック・ヘヴィメタル」、「ヒップホップ」、「ダンス・エレクトロニカ」、「ソウル・R&B」、「ジャズ」、「クラシック」、「ニューエイジ・イージーリスニング」、「サウンドトラック」、「アニメ・ゲーム」)^{*12, *13}を参照し、これらのジャンルをカバーするように VOCALOID 楽曲動画を視聴した上で、楽曲動画に付与されたタグを参照する。次に、視聴した楽曲動画に適合するジャンルが 12 項目の中にあれば、そのジャンルをラベルとして選択する。異なるラベルを選択した楽曲動画について、タグの傾向が似ている場合は、1 つのラベルとしてまとめる。同じジャンルラベルを選択した楽曲動画について、タグの傾向が異なる場合には、そのジャンルを分割し、別々のラベルを選択する。視聴した楽曲動画にふさわしいジャンルラベルがない場合は、適切なジャンルラベルを新しく提案する。

b ジャンルラベルに対応するタグの選定

(a) で提案したラベルにタグを対応付ける。タグは、VOCALOID 関連のタグ一覧^{*14}の記事に

*12 http://www.amazon.co.jp/MP3-ダウンロード-音楽配信-DRMフリー/b/ref=topnav_storetab_dmusic?ie=UTF8&node=2128134051 (2014 年 7 月参照)

*13 「デジタルミュージック」のカテゴリを使用したのは、VOCALOID 楽曲動画が、CD という媒体ではなく MP3 で配信されるものであるため、より本研究で扱うジャンル分類に適合するという考えによる。

*14 <http://dic.nicovideo.jp/a/vocaloid> 関連のタグ一覧

記載されている、「ジャンルによる分類」タグ 82 個を利用する。

(2) 実験参加者間の一致件数を調査し、一致人数を考慮してジャンルラベルと対応するタグを決定

手順 (1) で選定したジャンルラベルならびに対応するタグについての、実験参加者間の一致件数を表 2、表 3 に示す。

表 2 選定したジャンルラベルの実験参加者間の一致件数

案の一致人数	件数	割合
3 人	4	11.8%
2 人	14	41.2%
1 人	16	47.0%

表 3 ジャンルラベルに対応するタグの実験参加者間の一致件数

案の一致人数	件数	割合
3 人	33	27.3%
2 人	40	33.1%
1 人	48	39.6%

実験参加者間で一致したラベルと対応するタグは優先して選定し、案が一致しない場合は、以下の方針で対応する。

- ラベルの名称が異なるが、その下に分類されたタグは、ほぼ共通する場合は、分類されたタグとの関係が、より明確に分かるラベルを選択する。
- ラベルに対するタグの分類が異なる場合は、タグに関するニコニコ大百科^{*15}の定義等を参照し、客観的な視点で選択する。

(3) クリエータのジャンル分類の実現可能性の検証

3.1 節で収集した、楽曲動画とタグのデータを使用し、1 人のクリエイターが投稿した楽曲動画集合に付与されたタグの中の「ジャンルによる分類」^{*16}タグと、手順 (2) で選定したタグを照合し、対応するジャンルラベルをクリエイターごとに付与する。その後、ジャンルラベルが付与されたクリエイターの割合や、クリエイター 1 人あたりに付与されたジャンルラベルの数を調査する。

§3 結果

選定した 21 件のジャンルラベルと、それに対応するジャンルタグのリストを表 4 に示す。

また、ジャンルラベルが付与されたクリエイターの割合と、クリエイター 1 人あたりに付与されたジャンルラベルの数を表 5 に示す。

§4 考察

各実験参加者は、参考として提示した Amazon.co.jp のジャンルラベルの中から、「ポップス」、「ロック」、「ハードロック・ヘヴィメタル」、「ヒップホップ」、「ダンス・エレクトロニカ」、「ソウル・R&B」、「ジャズ」、「クラシッ

表 4 ジャンルラベルと対応するタグのリスト

ラベル	タグ
ポップス	ボカロ歌謡曲, MikuPOP, VOCAPOP
テクノポップ	ミクノポップ, ミキノポップ
バラード	ボカロバラード, アコギミク, 初音ミクの弾き語り, ボカロアコギ
ロック	ボカロガレージ, VOCAROCK, VOCALEAMO, ミクリーモ, メグリーモ, VOCALLOUD, VOCALOID デジタルロック, オルタナティブミック, プログロイド, プログレミック, プログレン, ボカログランジ, ミクゲイザー, ボカロゲイザー, リンダストリアル, ボカロインドストリアル, ボカロロックバラード, 昔の V 系っぽいボカロ曲リンク, ポストミック, ボカロマンチェ
パンク	VOCAPUNK, ミクパンク, ミクウェーブ
ハードロック・ヘヴィメタル	ボカロメタル, ミクメタル, メタルカ, ミク HR/HM, ミクグランド
ヒップホップ	ボカロラップ, ボカラップ, ミックホップ, クラブで使えるボカロ曲リンク
ダンス	VOCALOID-EUROBEAT, DubSteloid, ボカ ロ ワ ル ツ, VoCASamba, ボカロサンバ, ボカロタンゴ
エレクトロニカ	ミクトロニカ, エレクトロルカ, ボカトロニカ, ミクビエント, ボカンビエント, ミクマル, VOCANOISE, ミクノイズ
ソウル・R&B	ボカソウル, VocaR&B, ボカロファンク, ボカロフュージョン, VOCASKA, ボカロブルース, ボカロレゲエ
ジャズ	VOCAJAZZ
クラシック	ボカロクラシカ, ミクラシック, ボカロシンフォニー, ボカロ吹奏楽, ボカロ聖歌隊, ボカロ古楽, ボカロワルツ
和風	VOCALOID 和風曲, ボカロ音頭, 盆踊ロイド, ボカロ演歌
民族調	VOCALOID 民族調曲
中華曲	VOCALOID 中華曲, 中華ロイド
カントリー	ボカロ民謡, ボカロトラッド, ボカントリー, 牧歌ロイド, ボカロ童謡, ボカロ尋常小學生, ボカノバ, ミクボッサ, VoCASamba, ボカロサンバ, ボカロタンゴ, VOCALATIN
洋楽	巡音洋楽シリーズ, 初音洋楽シリーズ, 鏡音洋楽シリーズ, ボカロ洋楽カバー
ハウス・トランス	ボカロハウス, MikuHouse, ルカハウス, ボカロトランス, ミクトランス, ルカトランス, サイケデミックトランス, ミックンベース, ボカムンベース, ミクノ, ボカロテクノ
現代音楽	VOCALOID 現代音楽, ボカロ渡谷系
ピアノ	ボカロピアノ, ピアノミク, ピアノルカ, ピアノリン, ピアノレン, ピアノ MEIKO, ピアノ GUMI
電波系	VOCALOID 電波ソング, Chiptune × VOCALOID リンク

ク」を選択していた。一方、実験参加者のコメントによると、「J-POP」は、「ポップス」との区別が難しいことから、選択されなかった。また、「ニューエイジ・イージーリスニング」、「サウンドトラック」、「アニメ・ゲーム」は、VOCALOID を使用したオリジナル楽曲動画の中に該当するものが存在しないため、選択されなかった。

表 5 ジャンルラベルが付与されたクリエイターの割合とクリエイター 1 人あたりに付与されたジャンルラベルの数

ジャンルラベルの数	クリエイター数	割合
0	236	18.37 %
1	290	22.57 %
2	256	19.92 %
3	167	13.00 %
4	114	8.87 %
5	105	8.17 %
6	38	2.96 %
7	40	3.11 %
8	18	1.40 %
9	9	0.70 %
10	6	0.47 %
11	2	0.16 %
12	3	0.23 %
13	1	0.08 %
総数	1,285	100.00 %

*15 <http://dic.nicovideo.jp/>

*16 <http://dic.nicovideo.jp/a/vocaloid> 関連のタグ一覧#genre

この他に新たに定義した、「バラード」、「テクノポップ」、「パンク」、「和風」、「民族調」、「中華曲」、「カントリ」、「洋楽」、「ハウス・トランス」、「現代音楽」、「ピアノ」、「電波系」のラベルについては、「ジャンルに関する分類」タグにその特徴を表すものが存在するため、作成されたと考えられる。特に、「和風」や「民族調」の曲は、VOCALOID 楽曲動画に多く見られるため、ジャンルの分類に有効と考える。

また、表 5 から、1,285 名のクリエイターのうち、1,049 名のクリエイターに関して、1~13 件のジャンルラベルが付与された。ラベルが付与されなかった 236 名に関しては、投稿した楽曲動画にジャンル分類タグが付与されていない。この結果から、表 4 のジャンル分類に基づいて、約 82% のクリエイターを音楽ジャンルによって分類できることを確認した。なお、ジャンルが付与されなかった約 18% のクリエイターは、5.2 節で説明する重み付けについて、音楽ジャンルに関する重みを加算せずに、VOCALOID 名と楽曲動画の印象とを手がかりとして利用することで検索する。

4.2 評価実験：楽曲動画の印象の自動分類

§1 目的

クリエイターの制作している楽曲動画に対する印象の自動分類の精度を評価し、クリエイターの検索の手がかりとして利用できるか検証する。

§2 実験データ

印象の分類には、山本ら [山本 13a] によって構築された、楽曲動画印象データセットを利用する。このデータセットは、「ニコニコ動画」に投稿された楽曲動画のうち、「VOCALOID」タグの付与された、2012 年 8 月時点で再生数の多い楽曲動画上位 500 件に対して、印象の評価を行ったものである。印象クラスには「堂々」、「元気が出る」、「切ない」、「激しい」、「滑稽」、「可愛い」、「Valence」、「Arousal」の 8 つのクラスが存在する。山本ら [山本 13a] は、MIREX [Hu 08] で使用された 5 つの印象に、ニコニコ動画上で多く存在する「可愛い」を加えて、6 つの印象を定義した。さらに、Russel [Russel 80] が提案した Valence-Arousal 空間を参考にして、2 つの印象クラスを追加している。Russel の定義では、Valence は快-不快を表し、Arousal は覚醒-鎮静を表す。山本らは、既存研究を参考にした上で、印象クラスを表す語群として、Valence については、「明るい (暗い) 気持ちになる」、Arousal については、「強気な (弱気な)」などを用いている。データセットには、それぞれの印象クラスに対して、最低 3 名の評価者が 5 段階評価を行った評価の平均値が付与されている。

§3 実験方法

本節では、上記のデータセットに含まれる 8 つの印象クラスについて、LIBSVM を用いた 2 値分類器を作成し、交差検定を行い、その F 値、精度、再現率、正解率

を測る。

交差検定には、LIBSVM で用意されている easy.py を利用する。easy.py は、SVM のパラメータを調整し、デフォルトで 5 分割交差検定を行い、その正解率を測るプログラムである。SVM のカーネルは RBF を採用した。パラメータは、easy.py に備わっているグリッドサーチにより最適化した。

LIBSVM の入力データは、以下の手順で作成する。

- (1) まず、楽曲動画印象データセットに含まれている楽曲動画データ 500 件のうち、3.1 節で収集した楽曲動画データと一致する 272 件を抽出する。
- (2) 次に、それぞれの印象に関して、以下のように楽曲動画データのラベルを設定する。
 - a 「堂々」、「元気が出る」、「切ない」、「激しい」、「滑稽」、「可愛い」に対する 1~5 の評価値のうち、3 より大きい値をとる楽曲動画を正例、3 より小さい値をとる楽曲動画を負例とする。
 - b 「Valence」、「Arousal」に対する -2~2 の評価値のうち、0 より大きい値をとる楽曲動画を正例、0 より小さい値をとる楽曲動画を負例とする。
- (3) 楽曲動画の視聴者コメントを利用して、印象分類の手がかりとする素性を抽出する。視聴者コメントのデータは 3.4 節で収集したものを使用し、ニコニコ大百科、はてなキーワード^{*17}、Wikipedia^{*18}の見出し語を登録して辞書を拡張した形態素解析器 MeCab^{*19}を用いて、形態素ごとにわかち書きを行う。印象分類に有効な素性を選択するため、それらの形態素と印象との相互情報量 (expected mutual information) [Manning 08] を計算する。なお、形態素の品詞は、幅広く手がかりを収集するために、特に制限は設けていない。相互情報量は、2 つの確率変数の相互依存の尺度を数値化したものである。印象と形態素についての相互情報量を求めることで、その印象に特徴的な形態素を求めることができる。本研究で用いる相互情報量は、印象と形態素ごとに、以下の式 (1) で定義する。

$$I(U; C) = \frac{N_{11}}{N} \log_2 \frac{N N_{11}}{N_{1*} N_{*1}} + \frac{N_{01}}{N} \log_2 \frac{N N_{01}}{N_{0*} N_{*1}} + \frac{N_{10}}{N} \log_2 \frac{N N_{10}}{N_{1*} N_{*0}} + \frac{N_{00}}{N} \log_2 \frac{N N_{00}}{N_{0*} N_{*0}} \quad (1)$$

ここで、 U は楽曲動画の視聴者コメントとして特定の形態素が出現するか否かを表す確率変数で、 C は楽曲動画が特定の印象に該当するか否かを表す確率変数である。 N_{11} は、その形態素をコメントに含み、その印象に該当する楽曲動画数を表す。同様に、 N_{10} は、その形態素をコメントに含むが、その印象に該当しない楽曲動画数を、 N_{01} は、その形態素をコメ

*17 <http://d.hatena.ne.jp/keyword/>

*18 <http://ja.wikipedia.org/wiki/メインページ>

*19 <http://mecab.googlecode.com/svn/trunk/mecab/doc/index.html>

ントに含まないが、その印象に該当する楽曲動画数を、 N_{00} は、その形態素をコメントに含まず、その印象に該当しない楽曲動画数を表す。また、 N_{1*} は、 $N_{11} + N_{10}$ のように、その形態素をコメントに含む楽曲動画数を表す。同様に、* を用いることで、 N_{0*} 、 N_{*1} 、 N_{*0} は、その形態素をコメントに含まない楽曲動画数、その印象に該当する楽曲動画数、その印象に該当しない楽曲動画数を表す。最後に、 N は楽曲動画の総数であり、 $N = N_{00} + N_{01} + N_{10} + N_{11}$ となる。

相互情報量のしきい値は、経験的に調整して 0.25 と設定した。また、92 語のストップワードリストと照合してストップワードを取り除いた。ストップワードリストの一部を以下に示す。

ストップワードリストの一部

あれ、それ、そんな、っ、っけ、でしょ、という、とても、どっか、なれる、るな、コレ、ネ、三、>、●

以上の手順により、印象ごとに選択された形態素の集合を素性として用いた。素性の値は、各動画についてのすべてのコメントに対する形態素の頻度とした。

§4 評価尺度

印象分類にあたっては、F 値、精度、再現率、正解率の 4 つの評価尺度を採用する。これらは、以下の式 (2)~(5) で定義する。

$$F \text{ 値} = \frac{2 \times \text{精度} \times \text{再現率}}{\text{精度} + \text{再現率}} \quad (2)$$

$$\text{精度} = \frac{TP}{TP + FP} \quad (3)$$

$$\text{再現率} = \frac{TP}{TP + FN} \quad (4)$$

$$\text{正解率} = \frac{TP + TN}{TP + FP + TN + FN} \quad (5)$$

なお、上記の式中の TP は正しく正例と判断したもの、FP は誤って正例と判断したもの、FN は誤って負例と判断したもの、TN は正しく負例と判断したものとす。

§5 結果

楽曲動画印象データセットから抽出した 272 件の楽曲動画に対して、8 つの印象ごとに 2 値分類器を作成し、5 分割交差検定を行った。評価結果を表 6 に示す。

6 つの印象の平均値は、マクロ平均である。太字は、各評価尺度において最も高い値を表している。

§6 考察

F 値に関してみると、8 つの印象クラスすべてに関して 0.85 以上となっており、特に、「可愛い」、「激しい」と「Valence (明るい/暗い)」、「Arousal (強気な/弱気な)」の判別は 0.9 を超えている。また、再現率を見ると、「滑稽」という印象の分類の再現率が 0.75 に満たない。これは、データの正例数がデータ数の 15% と少なく、偏りがあつたためと思われる。これらの印象が、クリエイタ検

表 6 各印象に関する 2 値分類器の F 値、精度、再現率、正解率

印象	F 値	精度	再現率	正解率
堂々	0.853	0.838	0.869	0.766
元気が出る	0.884	0.941	0.833	0.838
切ない	0.868	0.958	0.793	0.774
激しい	0.911	0.935	0.888	0.781
滑稽	0.853	1.000	0.744	0.916
可愛い	0.931	0.964	0.900	0.874
6 つの印象の平均値	0.883	0.939	0.838	0.825
Valence	0.903	0.851	0.961	0.791
Arousal	0.922	0.874	0.975	0.831

索の手がかりとして用いることができるかどうかについては、5 章の実験で検証する。

5. 提案：クリエイタ検索のためのランキング

本章では、クリエイタ検索のためのランキング手法について説明する。クリエイタに楽曲動画の制作を依頼するにあたっては、意図した楽曲動画を過去に制作しているかという点が鍵となる。以上を踏まえ、クリエイタの検索の適合性は、そのクリエイタが過去に作成した適合動画数によって判断する。また、提案するランキング手法の妥当性を検証するため、サンプリング調査を行う。

5.1 クリエータ検索のためのランキング手法

- (1) 3.1 節で収集した 22,593 件の楽曲動画のデータに対して、クリエイタの各楽曲動画について VOCALOID 名を付与する。
- (2) 4.1 節で定義したジャンル体系に基づき、クリエイタの各楽曲動画について音楽ジャンルを付与する。
- (3) 4.2 節で作成した分類器を用いて、クリエイタの各楽曲動画について印象を付与する。印象ラベルは、「堂々」、「元気が出る」、「切ない」、「激しい」、「滑稽」、「可愛い」の 6 つと、「Valence」に対応する「明るい」、「暗い」、および「Arousal」に対応する「強気な」、「弱気な」の 4 つを合わせた 10 個を用いる。なお、本研究では、最初の 6 つの印象がクリエイタ検索の手がかりとして利用できない場合に、「Valence」を、主にリズムや音響から受ける印象の楽曲に対する手がかりとして利用し、「Arousal」は、主に歌詞から受ける印象の楽曲に対する手がかりとして用いる。
- (4) データを整形して MySQL^{*20} のデータベースに格納し、検索課題に合致したものを検索結果として表示する。なお、検索課題のキーは、手動で判断して入力した。データベースの構造を表 7 に示す。
- (5) 結果を表示する際、ランキングを行う。ランキングに利用するスコアを、VOCALOID 名、音楽ジャンル、楽曲動画の印象の 3 つの属性のうち、少なくともどれか 1 つに適合する楽曲動画を 5 件以上制作

*20 <http://www.jp.mysql.com/>

表 7 検索用データベースの構造

列数	内容
1	動画 ID (キー)
2	クリエイタ名
3~53	51 件の VOCALOID 名に対する 1 か 0 の値
54~74	21 件の音楽ジャンルに対する 1 か 0 の値
75~84	10 件の楽曲動画の印象に対する 1 か 0 の値

しているクリエイタに対して付与する。すなわち、制作した楽曲動画数が 5 件未満である 294 名のクリエイタについては、検索対象外とする。

楽曲動画の適合性スコアおよびクリエイタのスコアは、以下の式 (6)、式 (7) によって算出する。

$$\text{スコア}_{\text{楽曲動画の適合性}} = (b_{\text{vocaloid 名}} \times \alpha + 1) \times (b_{\text{音楽ジャンル}} \times \beta + 1) \quad (6)$$

$$\text{スコア}_{\text{クリエイタ}} = \frac{\sum \text{スコア}_{\text{楽曲動画の適合性}}}{\text{制作楽曲動画総数}} \times (b_{\text{楽曲動画の印象}} \times \gamma + 1) \quad (7)$$

なお、式 (7) におけるスコア_{クリエイタ} は、クリエイタごとに付与するスコアであり、楽曲動画の適合性を表すスコア_{楽曲動画の適合性} を、クリエイタが制作した楽曲動画集合について合計し、クリエイタの制作した楽曲動画の総数で割ることにより正規化を行う。スコアを正規化しているのは、正規化を行わないと、投稿した楽曲動画数の多いクリエイタが、特定の検索属性のみが適合する楽曲動画が多くなる場合に、高いスコアが与えられて、ランキングの上位を占めてしまい、不適切なランキングとなることによる。

$b_{\text{検索属性}}$ は、楽曲動画に付与された検索属性 (VOCALOID 名, 音楽ジャンル, 楽曲動画の印象) の値が検索課題の検索属性の値と一致していれば 1, そうでなければ 0 の値をとる。 $b_{\text{vocaloid 名}}$, $b_{\text{音楽ジャンル}}$, $b_{\text{楽曲動画の印象}}$ のうちどれか 1 つでも 0 件であった場合にスコアが 0 として算出されることを考慮して、それぞれの値に 1 を加算する。 α , β , γ は 3 つの属性に付与する重みである。

5.2 ランキング手法の妥当性の検証

§1 目的

サンプリング調査により、楽曲動画に付与された各検索属性の推定精度を評価し、クリエイタ検索の手がかりとして妥当かどうか検証する。

§2 調査方法

(1) 検索課題の作成

VOCALOID 楽曲動画をよく視聴する課題作成者 3 名に依頼し、VOCALOID 名, 音楽ジャンル, 楽曲動画の印象の 3 つの検索属性を手がかりとして探したいクリエイタを、検索課題として提案させた。なお、検索属性の具体例は示さずに、自由に課題を作

成させた。課題作成者 3 名の年齢, 性別, 職業は以下の通りである。

ア 20 代前半, 女性, 大学生

イ 20 代後半, 女性, 会社員

ウ 20 代前半, 男性, 家事手伝い

その結果, それぞれ 16 件, 13 件, 10 件の検索課題が提示された。本調査では, この 39 件の検索課題のうち, 各検索属性の推定精度を評価するためのサンプルとして, 課題作成者 1 人あたり 4 件ずつ (合計 12 件) の検索課題を選択した。検索課題の一覧を, 表 8 に示す。

表 8 サンプリング調査に使用した検索課題の一覧

検索課題	課題作成者
冰山キヨテルで激しいロック	ア
神威がくぼで暗いメタル	ア
初音ミクで激しいトランス	ア
初音ミクで元気が出るテクノポップ	ア
結月ゆかりで明るいジャズ	イ
開発コード miki で明るいテクノ	イ
鏡音リンで堂々としたダブステップ	イ
初音ミクで可愛いチップチューン	イ
巡音ルカで明るいエレクトロニカ	ウ
鏡音リンで格好良いスカ	ウ
MEIKO でおしゃれなジャズ	ウ
GUMI で可愛いユーロビート	ウ

(2) 適合性の評価

12 件の検索課題について 3 つの検索属性をすべて満たすと推定した楽曲動画を, 各課題ごとに 3 件ずつランダムに抽出し, 合計 36 件の楽曲動画について, 各検索属性の精度 (Precision) を調査するとともに, 第一著者による主観評価により検索課題に適合した楽曲動画を検索できていることを確認する。

§3 結果

36 件の楽曲動画に関する各検索属性の推定精度は, VOCALOID 名が 0.97, 音楽ジャンルが 0.89, 楽曲動画の印象が 0.81 となった。また, 主観評価により検索課題に適合すると判定された楽曲動画の推定精度は, 0.72 となった。

§4 考察

VOCALOID 名は, ほぼ正しく推定できていたものの, VOCALOID 名が付与された楽曲動画において, その VOCALOID が他の VOCALOID のコーラスを担当している楽曲動画が 1 件あった。このケースは, 楽曲動画の制作を依頼する観点から不適合と判定した。こうしたケースは見られるものの, 推定精度が高いことから, VOCALOID 名は, クリエータ検索の手がかりとして利用できる考える。

また, 音楽ジャンルや楽曲動画の印象は, 8~9 割の動画に対して適切な自動推定を行っており, 同様にクリエイタ検索の手がかりとして利用できる考える。

6. クリエータ検索システムの評価

本章では、クリエイータ検索手法の評価について述べる。

6.1 目的

以下の3つの目的で比較実験を行う。

- (1) 既存のジャンルとの比較により、提案する音楽ジャンルの有用性を明らかにする。既存のジャンルには、4.1節で利用した、Amazon.co.jpの「デジタルミュージック」における「ジャンル別ベストセラー」の12項目を用いる。
- (2) 音楽ジャンル、楽曲動画の印象の2つの検索属性の効果を検証するために、各検索属性を手がかりから除いた手法との比較を行い、それぞれの検索属性を利用してクリエイータを検索することが有効な課題について考察する。
- (3) 提案手法の各検索属性が検索にどの程度貢献しているのかを明らかにする。

6.2 実験方法

§1 評価に使用する検索課題

5.2節で作成した39件の検索課題のうち、サンプリング調査に使用した検索課題とは異なる検索課題を、課題作成者1人あたり4件ずつ(合計12件)、本比較実験のために選択した。検索課題の一覧を、表9に示す。

表9 比較実験に使用した検索課題の一覧

課題ID	検索課題	課題作成者
A	巡音ルカで切ないジャズ	ア
B	鏡音リンで可愛いR&B	ア
C	KAITOでおしゃれなR&B	ア
D	MEIKOで前向きなバラード	ア
E	鏡音リンで爽やかなトランス	イ
F	開発コード miki で可愛いテクノポップ	イ
G	神威がくぼで強気なロック	イ
H	鏡音レンで元気が出るヒップホップ	イ
I	初音ミクで面白げなハウス	ウ
J	鏡音レンで激しいロック	ウ
K	GUMIでおしゃれなエレクトロニカ	ウ
L	IAで力強いバラード	ウ

§2 楽曲動画のスコアに対するパラメータの値の調整

表8の検索課題を利用して、式(6)における各検索属性の重みを1~10の範囲で最適化し、それぞれ $\alpha = 7$, $\beta = 4$, $\gamma = 3$ とした。

§3 評価方法

検索課題に従って検索を行い、結果として表示されたランキングのうち、上位10件について4段階で適合性を判定する。適合性判定は、第一著者が行い、課題作成者に評価の妥当性を確認した。適合性の評価値は、ランキングされたクリエイータに対してユーザが楽曲の制作を

依頼する際に、検索課題により意図した楽曲動画を制作する能力があるかという点を判断する基準として、適合する楽曲動画の件数を用いて、以下のように設定する。

- 3点 検索課題に合っている楽曲動画を5件以上制作している
- 2点 検索課題に合っている楽曲動画を2~4件制作している
- 1点 検索課題に合っている楽曲動画を1件制作している
- 0点 検索課題に合っている楽曲動画を制作していない

評価値を用いて、nDCG (normalized Discounted Cumulative Gain) [Järvelin 02] および精度を算出する。nDCG、精度は、 $p = 10$ として、以下の式(8), (10)で定義する。

$$nDCG_p = \frac{DCG_p}{IDCG_p} \quad (8)$$

$$DCG_p = rel_1 + \sum_{i=2}^p \frac{rel_i}{\log_2 i} \quad (9)$$

$$精度 = \frac{|A \cap B|}{|B|} \quad (10)$$

なお、上記の式(8)中の $IDCG_p$ は、理想的なランキング(すべてのランク n に対して、評価値 $_n \geq$ 評価値 $_{n+1}$)の時のDCGである。また、式(10)中の B は提案手法によって検索されたクリエイータの集合であり、 A は適合性判定において2点、3点を付与されたクリエイータの集合とする。すなわち、精度は、検索課題に適合している楽曲動画を2件以上投稿しているクリエイータの割合を考慮して評価する尺度として用いる。

nDCGは、ランキングの上位10件において、評価値の高いクリエイータが評価値の低いクリエイータよりも高い順位である場合に値が大きくなる。一方、精度は、ランキングの上位10件において、検索結果として適切なクリエイータがどの程度含まれているかという割合を示す。

§4 比較手法

提案手法による検索との比較手法として、以下の3つの手法で検索を行う。

提案手法 VOCALOID名と音楽ジャンルと楽曲動画の印象による検索

比較手法1 VOCALOID名と既存の音楽ジャンル(Amazonのジャンル)と楽曲動画の印象による検索

比較手法2 VOCALOID名と音楽ジャンルによる検索(楽曲動画の印象は用いない)

比較手法3 VOCALOID名と楽曲動画の印象による検索(音楽ジャンルは用いない)

なお、比較手法1は、Amazonのジャンル12項目に対応するよう、提案する音楽ジャンルのラベルを、以下のように変更した。

- ハウス・トランス \Rightarrow ハウスをソウル・R&Bに、トランスをダンス・エレクトロニカに分割
- ダンスとエレクトロニカ \Rightarrow ダンス・エレクトロニカにまとめる

- テクノポップ ⇒ ポップスに融合
- バラード ⇒ ポップスに融合
- カントリー ⇒ ポップスに融合
- パンク ⇒ ロックに融合

また、以下に示す比較手法 4~6 についても評価を行い、各検索属性が、どの程度貢献しているのかを明らかにする。

比較手法 4 VOCALOID 名のみを用いた検索

比較手法 5 音楽ジャンルのみを用いた検索

比較手法 6 楽曲動画の印象のみを用いた検索

6.3 結果

提案手法による検索の nDCG および精度の評価値と、各比較手法による検索の nDCG および精度の評価値を、表 10、表 11 に示す。

表 10 提案手法および比較手法 1, 2, 3 についての評価結果

課題 ID	提案手法 VOCALOID 名 +音楽ジャンル +楽曲動画の印象		比較手法 1 VOCALOID 名 +Amazon ジャンル +楽曲動画の印象		比較手法 2 VOCALOID 名 +音楽ジャンル		比較手法 3 VOCALOID 名 +楽曲動画の印象	
	nDCG	精度	nDCG	精度	nDCG	精度	nDCG	精度
A	0.33	0.30	0.33	0.30	0.31	0.30	0.26	0.20
B	0.18	0.20	0.14	0.20	0.08	0.10	0.14	0.20
C	0.28	0.20	0.14	0.20	0.30	0.20	0.06	0.10
D	0.72	0.70	0.56	0.70	0.68	0.60	0.64	0.70
E	0.49	0.40	0.18	0.20	0.30	0.20	0.29	0.20
F	0.38	0.10	0.31	0.10	0.37	0.10	0.38	0.20
G	0.88	0.90	0.88	0.90	0.85	0.90	0.71	0.70
H	0.30	0.20	0.30	0.20	0.41	0.20	0.09	0.10
I	0.61	0.70	0.32	0.40	0.69	0.80	0.05	0.10
J	0.94	1.00	0.90	0.80	0.88	0.80	0.60	0.60
K	0.37	0.20	0.34	0.20	0.32	0.20	0.07	0.00
L	0.39	0.40	0.37	0.30	0.43	0.40	0.28	0.20
平均値	0.49**	0.44*	0.40	0.38	0.47	0.40	0.30	0.28

***: 比較手法 1 に対して t 検定 (有意水準 5%, 両側検定) で有意に向上。
比較手法 3 に対して t 検定 (有意水準 1%, 両側検定) で有意に向上。
*: 比較手法 3 に対して t 検定 (有意水準 5%, 両側検定) で有意に向上。

表 11 比較手法 4, 5, 6 についての評価結果

課題 ID	比較手法 4 VOCALOID 名		比較手法 5 音楽ジャンル		比較手法 6 楽曲動画の印象	
	nDCG	精度	nDCG	精度	nDCG	精度
A	0.21	0.30	0.17	0.10	0.00	0.00
B	0.00	0.00	0.03	0.00	0.00	0.00
C	0.05	0.10	0.21	0.10	0.00	0.00
D	0.58	0.50	0.00	0.00	0.00	0.00
E	0.04	0.00	0.10	0.10	0.08	0.10
F	0.32	0.10	0.03	0.00	0.00	0.00
G	0.62	0.70	0.00	0.00	0.00	0.00
H	0.00	0.00	0.27	0.20	0.06	0.00
I	0.04	0.10	0.40	0.50	0.06	0.10
J	0.33	0.40	0.12	0.10	0.10	0.10
K	0.14	0.10	0.00	0.00	0.00	0.00
L	0.33	0.20	0.21	0.20	0.00	0.00
平均値	0.22	0.21	0.13	0.11	0.03	0.03

太字は、各手法ごとに算出された nDCG および精度において最も高い値を示している。

提案手法と比較手法 1 (Amazon の音楽ジャンルを考慮した場合) の nDCG については、t 検定 (有意水準 5%, 両側検定) により、有意差が確認できた。提案手法と比較手法 3 (音楽ジャンルを考慮しない場合) の nDCG に

ついては、t 検定 (有意水準 1%, 両側検定) により、有意差が確認できた。また、提案手法と比較手法 3 (音楽ジャンルを考慮しない場合) の精度については、t 検定 (有意水準 5%, 両側検定) により、有意差が確認できた。

一方、提案手法と比較手法 2 (楽曲動画の印象を考慮しない場合) については、有意差は確認できなかった。このことから、楽曲動画の印象については、検索課題に応じて有効な場合とそうでない場合とがあると考えられる。次の考察では、検索課題ごとの結果の差異について、議論をまとめる。

6.4 考察

提案手法と比較手法 1 (Amazon の音楽ジャンルを考慮した場合) との比較により、提案する音楽ジャンルの有用性について考察する。12 課題のうち、nDCG が向上したのは B, C, D, E, F, I, J, K, L の 9 課題、精度が向上したのは課題 E, I, J, L の 4 課題だった。

● 課題 E, I

課題 E, I は、評価値が大きく向上した。よって、「ハウス」と「トランス」は、1 つのジャンルとしてまとめることにより、適切なランキングを提供できる。

● 課題 B, C

課題 B, C は、nDCG がそれぞれ向上した。よって、「ソウル・R&B」から「ハウス」を別のジャンルとして分割することで、適切なランキングを提供できる。

● 課題 D, L

課題 D は、nDCG が大きく向上した。また、課題 L についても nDCG に向上が見られる。よって、「ポップス」から「バラード」を別のジャンルとして分割することで、適切なランキングを提供できる。

● その他の課題

課題 J と課題 K についても、nDCG はわずかではあるが向上した。この結果から、「ロック」から「パンク」を別のジャンルとして分割し、「ダンス・エレクトロニカ」から「エレクトロニカ」を別のジャンルとして分割することで、適切なランキングを提供できる可能性があると考えられる。

以上の考察から、提案する音楽ジャンルを利用することにより、既存のジャンルを利用する場合と比較して、適切なクエリを検索できることが明らかとなった。

次に、比較手法 2 (楽曲動画の印象を考慮しない場合) と比較手法 3 (音楽ジャンルを考慮しない場合) の評価値に対する提案手法の評価値の観点で、課題および音楽ジャンルと楽曲動画の印象について考察する。

(1) 音楽ジャンルにより評価値が向上した課題:

A, B, C, D, E, G, H, I, J, K, L

課題 A, B, C, D, E, G, I, J, K, L は、提案手法が比較手法 3 より評価値が向上しており、音楽

ジャンルによる検索が有効と考える。これらの検索に用いたジャンルは「ロック」、「ハウス・トランス」、「エレクトロニカ」、「R&B」、「バラード」、「ジャズ」、「ヒップホップ」である。これら7つのジャンルは、VOCALOID 楽曲における音楽ジャンルとして判別しやすく、ジャンルタグの付与も適切に行われていた。

特に、「ロック」(課題 G, J) と「ハウス・トランス」(課題 E, I)、「バラード」(課題 D, L) は、精度が高い。これらの課題では、上位にランキングされたクリエイターが多く、適合する楽曲動画を制作していた。したがって、このジャンルにこだわって楽曲動画を制作するクリエイターを多く発見できたと考えられる。一方、「エレクトロニカ」(課題 K) や「R&B」(課題 B, C) に関して見つかったクリエイターは、多岐にわたるジャンルを作れる人物が多く、一人一人の制作する適合する楽曲動画数自体は少数であった。しかし適合する楽曲動画には適切にタグ付けがされているものが多かったため、適合するクリエイターをうまく検索することができたと考えられる。また、課題 H は、印象を手がかりとしないほうが nDCG が向上する。これは、「ヒップホップ」というジャンルの楽曲動画自体が少ないために、「元気が出る」という印象により、過剰な絞り込みが行われたと考えられる。

(2) 楽曲動画の印象により評価値が向上した課題 : B, D, E, J

課題 B, D, E, J は、提案手法が比較手法 2 より評価値が向上しており、楽曲動画の印象による検索が有効と考える。特に B, E については nDCG に大きく向上がみられる。このことから、ジャンルによる検索では十分に絞り込みができていないと考えられ、楽曲動画の印象を検索属性とすることの妥当性を示唆している。

これらの検索に用いた印象は「可愛い」、「激しい」、「爽やかな」、「前向きな」というもので、「爽やかな」、「前向きな」という印象はそれぞれ「元気が出る」、「明るい」という印象で代用して検索を行った。これらの印象は、表 6 の F 値も比較的高く、検索課題で要求されている楽曲動画の印象についての揺れも少ないと考えられる。

(3) 評価値が著しく向上しなかった課題 : F

課題 F は、ジャンルを手がかりとすることで、精度が悪化している。これは、視聴者が「テクノポップ」と類似する音楽ジャンルである「テクノ」を区別できず、「テクノポップ」に対応するタグを、誤って付与した楽曲動画が、上位にランキングされたことによる。

最後に、比較手法 4 (VOCALOID 名による検索)、比較手法 5 (音楽ジャンルによる検索)、比較手法 6 (楽曲動画の印象による検索) の比較により、各検索属性の有

用性について考察する。半数以上の検索課題は、比較手法 4 の評価値が高い。検索課題において、VOCALOID 名と相性の良い音楽ジャンルや楽曲動画の印象を指定している場合には、VOCALOID 名が検索の手がかりとして有効である。一方で、課題 E, H, I は、比較手法 5 における精度が上回っている。これらの検索課題で使用されている VOCALOID 名は、様々な音楽ジャンルを対象とし、様々な印象を与える楽曲動画が制作されており、検索課題で指定した音楽ジャンルや楽曲動画の印象に適合する楽曲動画が相対的に少ない。その結果、「ヒップホップ」、「ハウス」など、音楽ジャンルを手がかりとした方が、適合する楽曲動画を検索しやすくなっている。

7. む す び

本稿では、ニコニコ動画に投稿された VOCALOID 楽曲動画に付与されたタグと視聴者コメントを利用して、VOCALOID を使用した楽曲動画を制作かつ発表しているクリエイターを、視聴者の好みに合わせて検索する手法を提案した。クリエイターを探す手がかりとしては、クリエイターが使用している VOCALOID のキャラクター名、制作している楽曲動画の音楽ジャンル、楽曲動画の印象の3つを用いた。

提案手法の実現のために、一般的な音楽ジャンルと、VOCALOID 楽曲動画に付与されたジャンルに関連したタグとを考慮することにより、クリエイターを検索するための 21 の音楽のジャンル体系を定義した。

また、楽曲動画の印象推定は、楽曲動画に付与された視聴者コメントを形態素解析し、各印象との相互情報量が高い形態素を素性として選択した上で、RBF カーネルを用いた SVM のパラメータを、グリッドサーチにより最適化した。8つの印象クラスについて、5分割交差検定によって分類精度を評価したところ、すべての印象クラスに関して F 値が 0.85 以上となった。

12 件の検索課題に対して、提案手法による検索を行い、各課題について上位にランキングされたクリエイター 10 名を評価したところ、精度がマクロ平均で 0.44 となり、検索課題に適合するクリエイターを平均して約 4 名発見することがわかった。この結果から、提案する検索手法により、ユーザが好みのクリエイターを発見することや、楽曲動画のクリエイターにコンテンツ制作の仕事を依頼できる可能性を明らかにした。

また、VOCALOID クリエータの検索のための音楽ジャンルを提案し、既存のジャンルとの比較実験を行った。提案する音楽ジャンルは、既存のジャンルを用いて検索した場合と比較して、nDCG が有意に向上した。この結果から、提案する音楽ジャンルが VOCALOID クリエータのランキングにおいて有用であることを明らかにした。

さらに、提案する検索手法と、音楽ジャンルまたは楽曲動画の印象を考慮しない場合との比較実験を行い、検

素課題によってどのような音楽ジャンルおよび楽曲動画の印象が有効であるかを考察した。音楽ジャンルは、ほとんどの場合、クリエイタの検索に対して有効だった。実験した範囲では、有効でないケースは、視聴者がタグを付与する上で、類似する音楽ジャンルと適切に区別できず、誤ったタグを付与している場合に限られた。楽曲動画の印象は、楽曲動画の印象が適切に推定された場合に、有効である傾向が見られた。

今後は、本システムを Web 上のサービスとして実現し、楽曲動画の制作を依頼するユーザが満足するかたちでクリエイタを検索できるかどうか検証する予定である。

謝 辞

本研究の一部は、科学研究費補助金基盤研究 B (課題番号 16H02913)、筑波大学研究基盤支援プログラム (Bタイプ) の助成を受けて遂行された。

本研究の一部の実験では、国立情報学研究所が提供する「ニコニコデータセット」を利用した。ここに深く感謝する。

◇ 参 考 文 献 ◇

- [Chen 09] Chen, L., Wright, P., and Nejdl, W.: Improving Music Genre Classification Using Collaborative Tagging Data, in *Proc. of the 2nd ACM Int'l Conf. on Web Search and Data Mining (WSDM 2009)*, pp. 84–93, Barcelona, Spain (2009)
- [Correa 16] Correa, D. C. and Rodrigues, F. A.: A Survey on Symbolic Data-based Music Genre Classification, *Expert Systems with Applications*, Vol. 60, pp. 190–210 (2016)
- [Costa 12] Costa, Y. M. G., Oliveira, L. S., Koerich, A. L., Gouyon, F., and Martins, J. G.: Music Genre Classification Using LBP Textual Features, *Signal Processing*, Vol. 92, No. 11, pp. 2723–2737 (2012)
- [土橋 06] 土橋 佑亮, 片寄 晴弘: SOM を用いたベースラインからの音楽ジャンル解析, 情報処理学会 研究報告 音楽情報科学, pp. 31–36 (2006)
- [Fu 11] Fu, Z., Lu, G., Ting, K. M., and Zhang, D.: Music Classification via the Bag-of-features Approach, *Pattern Recognition Letters*, Vol. 32, No. 14, pp. 1768–1777 (2011)
- [後藤 12] 後藤 真孝: CGM の現在と未来: 初音ミク, ニコニコ動画, ピアプロの切り拓いた世界, 情報処理, Vol. 53, No. 5, pp. 466–471 (2012)
- [Hu 08] Hu, X., Downie, J. S., Laurier, C., Bay, M., and Ehmann, A. F.: The 2007 MIREX Audio Mood Classification Task: Lessons Learned, in *Proc. of the 9th Int'l Conf. on Music Information Retrieval*, pp. 462–467, Philadelphia, USA (2008)
- [Hu 10] Hu, X. and Downie, J. S.: When Lyrics Outperform Audio for Music Mood Classification: A Feature Analysis, in *Proc. of the 11th Int'l Society for Music Information Retrieval Conference (ISMIR 2010)*, pp. 619–624, Utrecht, Netherlands (2010)
- [Jamdar 15] Jamdar, A., Abraham, J., Khanna, K., and Dubey, R.: Emotion Analysis of Songs Based on Lyrical and Audio Features, *Int'l Journal of Artificial Intelligence & Applications (IJAA)*, Vol. 6, No. 3, pp. 35–50 (2015)
- [Järvelin 02] Järvelin, K. and Kekäläinen, J.: Cumulated Gain-Based Evaluation of IR Techniques, *ACM Trans. Inf. Syst.*, Vol. 20, No. 4, pp. 422–446 (2002)
- [金罎 13] 金罎 愛, 中野 倫靖, 後藤 真孝, 菊池 英明: ポピュラー音楽における歌声の印象評価語を自動推定するシステム, 情報処理学会 研究報告 音楽情報科学 (2013)
- [Kim 10] Kim, Y. E., Schmidt, E. M., Migneco, R., Morton, O. G., Richardson, P., Scott, J., Speck, J. A., and Turnbull, D.: Emotion Recognition: A State of the Art Review, in *Proc. of the 11th Int'l Society for Music Information Retrieval Conference (ISMIR 2010)*, pp. 255–266, Utrecht, Netherlands (2010)
- [熊本 02] 熊本 忠彦, 太田 公子: 印象に基づく楽曲検索: 検索ニーズに合った印象尺度の設計, 情報処理学会 研究報告 自然言語処理, pp. 35–40 (2002)
- [Manning 08] Manning, C. D., Raghavan, P., and Schütze, H.: *Introduction to Information Retrieval*, Cambridge University Press (2008)
- [Ogino 15] Ogino, A. and Yamashita, Y.: Emotion-Based Music Information Retrieval Using Lyrics, in *Computer Information Systems and Industrial Management*, Vol. 9339 of *Lecture Notes in Computer Science*, pp. 613–622, Springer (2015)
- [大野 16] 大野 直紀, 土屋 駿貴, 中村 聡史, 山本 岳洋: 独立した音楽と映像に対する印象評価からの音楽動画の印象推定手法, 第 8 回データ工学と情報マネジメントに関するフォーラム (DEIM2016), pp. E3–5 (2016)
- [Russel 80] Russel, J. A.: A Circumplex Model of Affect, *Journal of Personality and Social Psychology*, Vol. 39, No. 6, pp. 1161–1178 (1980)
- [Seki 13] Seki, Y. and Miyajima, K.: Finding Impressive Social Content Creators, in *Proc. of ACM SIGIR Conf. on Research and Development on Information Retrieval (SIGIR 2013)*, pp. 1041–1044, Dublin, Ireland (2013)
- [Soleymani 13] Soleymani, M., Caro, M. N., Schmidt, E. M., Sha, C.-Y., and Yang, Y.-H.: 1000 Songs for Emotional Analysis of Music, in *Proc. of the 2nd ACM Int'l Workshop on Crowdsourcing for Multimedia (CrowdMM 2013)*, pp. 1–6, Barcelona, Spain (2013)
- [Sturm 14] Sturm, B. L.: A Survey of Evaluation in Music Genre Recognition, in *Adaptive Multimedia Retrieval: Semantics, Context, and Adaptation*, Vol. 8382 of *Lecture Notes in Computer Science*, pp. 29–66, Springer (2014)
- [佃 11] 佃 洗撰, 中村 聡史, 田中 克己: 視聴者のコメントに基づく動画検索および推薦システムの提案, 第 19 回インタラクティブシステムとソフトウェアに関するワークショップ (WISS2011) 論文集, pp. 78–83 (2011)
- [Wang 09] Wang, F., Wang, X., Shao, B., Li, T., and Ogihara, M.: Tag Integrated Multi-label Music Style Classification with Hypergraph, in *Proc. of the 10th Int'l Society for Music Information Retrieval Conference (ISMIR 2009)*, pp. 363–368, Kobe, Japan (2009)
- [山本 13a] 山本 岳洋, 中村 聡史: 楽曲動画印象データセット作成とその分析, ARG 第 2 回 Web インテリジェンスとインタラクティブ研究会 (2013)
- [山本 13b] 山本 岳洋, 中村 聡史: 視聴者の時刻同期コメントを用いた楽曲動画の印象分類, 情報処理学会論文誌 データベース (TOD), Vol. 6, No. 3, pp. 61–72 (2013)
- [Yang 12] Yang, Y.-H. and Chen, H. H.: Machine Recognition of Music Emotion: A Review, *ACM Trans. on Intelligent Systems and Technology (TIST)*, Vol. 3, No. 3, pp. 40:1–40:30 (2012)

[担当委員: 奥 健太]

2016 年 5 月 9 日 受理

著 者 紹 介



末吉 優

2015 年 筑波大学 情報学群 知識情報・図書館学類卒業。



関 洋平(正会員)

2005 年総合研究大学院大学情報学専攻博士後課程修了。博士(情報学)。同年豊橋技術科学大学情報工学系助手。2008 年コロンビア大学客員研究員。2015 年筑波大学図書館情報メディア系准教授, 現在に至る。自然言語処理, 意見分析, 情報アクセスの研究に従事。ACM, ACL, 情報処理学会, 電子情報通信学会, 言語処理学会, 日本データベース学会, 人工知能学会各会員。