

商品レビューに基づく SOM を用いた音楽アーティスト分類手法

上原裕也[†] 北山大輔[†]

[†] 工学院大学情報学部

1 はじめに

近年インターネットの普及により、あらゆる種類の楽曲の検索が可能になっている。その結果、個人が検索して聴くことができる楽曲の種類が大幅に増加している。一方で、このように多種多様な楽曲に触れる機会が増えた反面、ユーザ毎の嗜好に合った楽曲を検索することが困難になるという問題がある。このような中で、無数と存在している楽曲の中から、ユーザの嗜好に合った楽曲を推薦するための研究が数多く進められており、楽曲の音響特徴量からユーザの嗜好に合った楽曲を推薦する技術 [1] などが研究されている。そこで本研究では、楽曲毎ではなく音楽アーティスト（以下「アーティスト」）毎に注目し、Amazon に投稿されたレビューを利用することで、多数のアーティストの中からユーザの嗜好に合ったアーティストを推薦する手法を提案する。ここで楽曲単位ではなくアーティスト単位の推薦としたのは、アーティスト単位の推薦とすることでもユーザの嗜好に合った楽曲を推薦することに繋がることと、特徴抽出に用いるレビュー量を楽曲単位で行う場合に比べ相対的に増加させ精度を高めることが期待できるという 2 つの理由からである。

本稿では、ユーザの嗜好に合ったアーティストを推薦することを目的とし、以下に取り組む。

- アーティストのアルバムに対するレビューから、評価に値する単語群を抽出する。
- 2 種類の SOM を生成し、それらの比較からユーザの嗜好に合うアーティストを推薦する。

2 レビュー内の単語の抽出・評価方法

アーティストに対する印象について、本研究では「一般的な印象」と「全作品を通して感じる印象」の 2 種類があると考え、一般的な印象とは、あるア

ティストのことをよく知らない人が感じる印象である。例としては、あるアーティストについて、TV で流れている曲を聞いたことがある程度の認識の人が抱く印象である。そのような印象を抱く楽曲を検索するため、各アーティストのシングル売上ランキングを調べ、その中でオリコンランキング 10 位以内にランクインしたことがある上位シングルを「一般的に知られている作品」とする。全作品を通して感じる印象とは、文字通りあるアーティストの全作品を聞いたことがあるような人（所謂そのアーティストのファン）が抱く印象である。よって、この印象に対する SOM を生成する際には、アーティストのすべての作品が対象となる。ただし本稿では、試行のため対象アーティストの作品のレビューを全て取得することは現実的ではないため、シングル枚数の上限をオリコンランキング上位から数えて 10 件、アルバム枚数を 5 件としている。

アーティストの作品の集合レビューに対し形態素解析を行い、アーティストの印象に関連する単語の候補として、レビューから形容詞・形容動詞を取得する。しかし、それだけでは「メロディがいい」等といったような名詞とセットで意味のある単語が「いい」としか抽出されなくなる。そのため、「いい」や「素晴らしい」といった名詞とセットで有用な単語候補となる形容詞を事前に辞書として準備しておきそれを用いて、その直前に出てくる名詞を抽出する。そして、取得した名詞・形容詞・形容動詞をアーティストの印象を推定する単語候補群とする。

取得した単語候補群をアーティストの印象と関係があるか否かを人手で確認し、7 割以上の人間が関係あると答えた単語を、「アーティストの印象を推定する単語」とする。ここで、一般的に知られている作品のレビューから取得したアーティストの印象を推定する単語群を I1、アーティスト全ての作品のレビューから取得したアーティストの印象を推定する単語を I2 とする。

3 SOM の生成と比較方法

I1, I2 から SOM を生成する。ここで、I1 から生成した SOM は「一般的な印象の SOM」、I2 から生成し

A Classification Method of Music Artists using SOM based on Item Reviews

Yuya UEHARA[†] and Daisuke KITAYAMA[†]

[†] Faculty of Information, Kogakuin University

160-0023, Shinjuku, Japan

j110020@ns.kogakuin.ac.jp, kitayama@cc.kogakuin.ac.jp

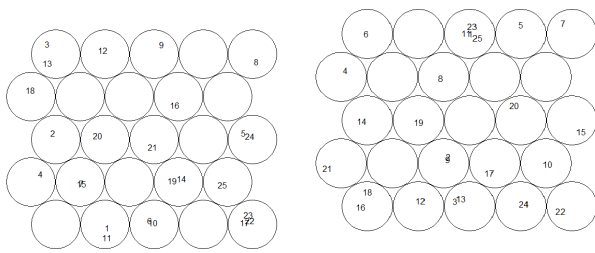


図 1: 一般的な印象の SOM
図 2: 全作品を通して感じる印象の SOM

た SOM は「全作品による印象の SOM」とする。それらの SOM を比較した際に、ユーザの嗜好に合っているアーティストから考えて推薦候補のアーティストの SOM 上の位置が (1) 一般的な印象と全作品を通して感じる印象がお互い近いもの、(2) 一般的な印象と全作品を通して感じる印象がお互い遠いもの、(3) 一般的な印象では近いが、全作品を通して感じる印象が遠いもの、(4) 一般的な印象では遠いが、全作品を通して感じる印象が近いもの、といった 4 種類に分類する。

この中で、(3)(4) に該当するアーティストをユーザの嗜好に合うと思われるアーティストとしてユーザに推薦する。それは、(3) の場合推薦対象のアーティストは元々知っている可能性が高いが、作品を通してみた印象が違ふと考えられる。よって、自分が気に入る楽曲が有りながらも、ユーザが聴いたことがないようなジャンルの楽曲に触れ得ることができるのではないかとと思われる。これによって、ユーザが新しいジャンルを自分の嗜好に合うジャンルとして受け入れ、趣味の幅を広げることができる。(4) の場合、推薦対象のアーティストは元々知らない可能性が高いが、作品を通した印象が似ていると思われる。つまり、ユーザからは一般的な印象から考えた際に聴かず嫌いをしている可能性があるが、実は作品をきちんと聴いてみるとユーザの嗜好に合う可能性が高いという推測から、ユーザに推薦する価値は高いと考えている。

4 評価実験と結果

今回は 5x5 のノードで SOM を生成し、最後にアーティストを格納する際の類似度の計算はユークリッド距離で行い、同じノード内、または隣接しているノード内に格納されているアーティストを距離が近いとすることとした。ユーザの嗜好に合っているアーティストとして「Janne Da Arc」と「B'z」を用意し、それぞれのアーティストの楽曲を好んで聴く 5 人に実験に協力してもらった。それぞれのアーティスト番号を 1,2

とし、推薦候補のアーティスト 23 種類を用意し、計 25 アーティストによる実験を行った。結果を、一般的なアーティストな印象の SOM が図 1、全作品を通して感じる印象の SOM が図 2 に示す。

ここからユーザに対して推薦対象となるアーティストは、前項の (3) 一般的な印象では近いが、全作品を通して感じる印象が遠いものに当てはまる 7 (GLAY), 15 (Mr.children) となる。(4) 一般的な印象では遠いが、全作品を通して感じる印象が近いものについては 1, 2 共に全作品を通して感じる印象の SOM 上で近い距離のアーティストが存在しなかった。

これらの結果を先ほどの協力者 5 人に、推薦されたアーティストが各自の嗜好に合うアーティストなのかどうかを「合う」か「合わない」の 2 択で評価してもらった。7 (GLAY) については 5 人中 5 人が、15 (Mr.children) については 5 人中 4 人が自分の嗜好に合っているアーティストであると回答した。

5 今後の課題

SOM の生成においてノードに格納する際にユークリッド距離を使用した。これではレビュー内の単語の抽出数によって差が出てしまうことが考えられる。よって、単語の抽出数を単語の総数で割ってから計算を行ったり、コサイン類似度を利用する等を行うことで改良が可能になると思われる。また本実験ではユーザの嗜好に合っているアーティストとして 2 つのアーティストを利用したが、この 2 つのアーティストが離れていたり、もっと複数のアーティストを利用した場合は推薦対象のアーティストが出力されないことが考えられる。よって、これの対策を考え改良することが今後の課題である。

謝辞

本研究の一部は、平成 25 年度科研費若手研究 (B) (課題番号: 24700098) によるものです。ここに記して謝意を表すものとします。

参考文献

[1] 荒川克憲, 小田川智, 松下文雄, 児玉泰輝, 塩田岳彦: 楽曲特徴量による嗜好音楽の解析, 情報科学技術フォーラム一般講演論文集 4(2), Vol16, No2 pp.275-276, 2006.